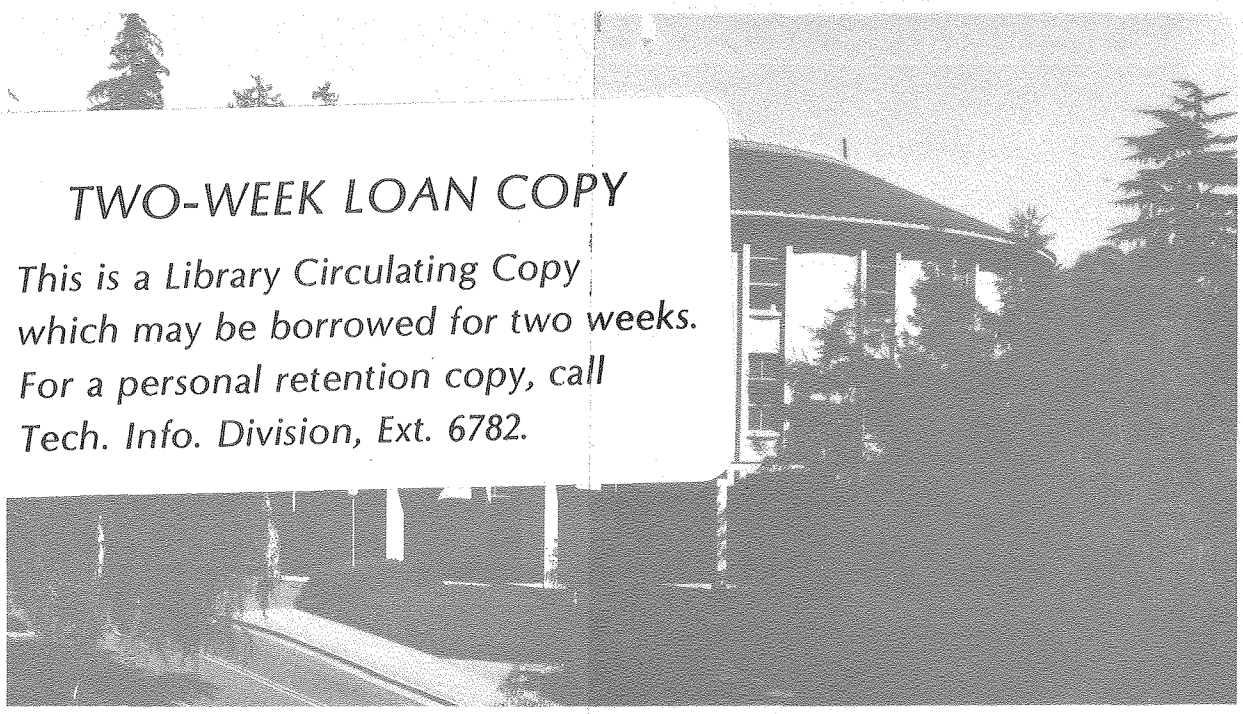
Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

CHEMICAL BIODYNAMICS DIVISION

I. A COMPARATIVE STUDY OF RIBO-, DEOXYRIBO-, AND HYBRID OLIGONUCLEOTIDE HELICES BY NUCLEAR MAGNETIC RESONANCE. II. OPTICAL STUDIES OF ETHIDIUM BINDING TO OLIGONUCLEOTIDES

Arthur Pardi
(Ph.D. thesis)

November 1980



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

DISCLAIMER

I. A COMPARATIVE STUDY OF RIBO-, DEOXYRIBO-, AND HYBRID OLIGONUCLEOTIDE HELICES BY NUCLEAR MAGNETIC RESONANCE

II. OPTICAL STUDIES OF ETHIDIUM BINDING TO OLIGONUCLEOTIDES

by

Arthur Pardi
Lawrence Berkeley Laboratory, University of California
Berkeley, CA 94720

ABSTRACT

The nonexchangeable base protons and the hydrogen bonding N-H-N imino protons were used to study the conformations and the helix-to-coil transitions in the following oligonucleotides: (I) a DNA duplex dCT_5G + dCA_5G; (II) an RNA duplex rCU_5G + rCA_5G; (III) a DNA-RNA hybrid duplex dCT_5G + rCA_5G; and (IV) a DNA-RNA hybrid triplex rCU_5G + dCA_5G. The first three mixtures all form stable double helical structures at 5°C, whereas IV forms a triple strand with a ratio of 2:1 rCU_5G:dCA_5G. The chemical shifts of the imino protons in the double strands indicate that I, II, and III have different conformations in solution. For example, the hydrogen bonded proton on one of the C·G base pairs is more shielded (a 0.4 ppm upfield shift) in helix I than in helix II or III. This implies a significant change in helical parameters, such as the winding angle, the distance between base pairs, or overlap of the bases. The coupling constants of the H1' sugar protons show that helix I has 90% 2' endo sugar conformation, whereas helix III has greater than 85% 3' endo conformation for the observed sugar rings. The chemical shift and sugar pucker data are consistent with helix I having

to my parents, for their support through the years, and for having the confidence in me to let me make my own decisions.

ACKNOWLEDGEMENTS

I would like to thank Nacho Tinoco for his help and guidance during the course of this work, and especially for giving me the independence to work out problems on my own and learn from my mistakes.

John Hearst provided valuable advice and help, especially in my early years. His boundless enthusiasm for his work and life was always a source of inspiration during hard times.

I wish to thank Melvin P. Klein for many fruitful discussions and for teaching me that there are many ways to approach a problem.

The whole Tinoco group provided an atmosphere where it was a pleasure to work and play. Barbara Dengler was always a source of good cheer, and also the expert on how to get things done around the lab or in the chemistry department. David Koh always seemed to be there when needed, whether as a source of oligonucleotides or just someone to complain to.

In my first year here Robert Lee helped me get my start in NMR while Charlie Reich was a source of friendly combativeness. I thank Bruce Johnson for many useful discussions (where he always had an answer for anything) and for his sense of humor and friendship.

The staff of Stanford Magnetic Resonance Laboratory (supported by NSF Grant GP 23633 and NIH grant RRoo711), especially Tony Ribeiro, provided useful advice and made my frequent trips to their lab much more pleasant.

The members of Mel Klein's research group provided much help and friendship while I was using their instrument. Greg Karczmar and Alan Koretsky were especially valuable in their help with instrumental problems and discussions on NMR.

Steve Winkle provided useful comments and advice, even though we did not always see eye-to-eye. Jeff Nelson took a lot of grief from me (some of it undeserved), and was still a willing source of valuable advice. David Keller helped with computer programming and more importantly helped improve my tennis game to its present (still pitiful) state. Joe Kao provided useful advice and critique. I wish to thank Carlos Bustamante for his willingness to help in whatever way he could and for his refreshing attitude on life.

Frank Martin provided innumerable invaluable discussions on the physical chemistry of nucleic acids, and was also the source of many of the oligonucleotides used in this work.

Kathy Morden took a lot of kidding that she didn't deserve, but always kept on smiling. I wish to thank her for countless hours in help on data collection on the NMR's. She also provided cookies and goodies for us to eat during the wee hours of the morning at Stanford, where the only way to keep awake is to keep eating (A. Pardi, unpublished result). Her collaboration provided a valuable source of ideas and help in sample preparation and data interpretation.

Ken Dahl was always willing to think about and work on any problem which I asked him to. He also provided valuable critiques of my work (and most other subjects) as well as numerous discussions.

The most valuable contribution all these people made was their friendship, which made this work enjoyable.

This work was supported by NIH Grant GM 10840 and by the Division of Biomedical and Environmental Research of the Department of Energy under Contract No. W-7405-ENG-48.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS..... i

TABLE OF CONTENTS..... v

PART I, A COMPARATIVE STUDY OF RIBO-, DEOXYRIBO-, AND HYBRID OLIGONUCLEOTIDE HELICES BY NUCLEAR MAGNETIC RESONANCE..... 1

CHAPTER I, INTRODUCTION..... 1

CHAPTER II, CONFORMATION OF OLIGONUCLEOTIDES STUDIED BY NMR..... 13

1. INTRODUCTION..... 13

2. MATERIAL AND METHODS..... 15

A) Oligonucleotides..... 15

B) Experimental NMR..... 16

3. RESULTS..... 20

A) Assignments of Nonexchangeable Protons..... 20

B) Assignment of the Base Paired Imino Protons.... 43

C) Temperature Dependence of the Nonexchangeable Protons in the Single Strands..... 54

D) The Helix-to-Coil Transition of $dCT_5G + dCA_5G$ and $dCT_5G + rCA_5G$ as Followed by the Chemical Shifts of the Base Protons..... 61

E) Temperature Dependence of the Exchangeable Imino Protons..... 75

F) Sugar Pucker of the Furanose Ring in the Double Strand..... 80

G) Linewidths of the Base Protons and Chemical Exchange..... 85

H) Base Pair Overlap in the Double Strand..... 93

4. DISCUSSION

A) Comparison of the Structures of Oligonucleotide Complexes in Solution..... 94

I) The Double Stranded Structures..... 94

II) The Triple Strand..... 99

B) Analysis of Thermodynamic Parameters by Chemical Shift Versus Temperature Plots of the Base Protons..... 101

C) Fraying Versus Melting of Oligonucleotides..... 109

5. CONCLUSION..... 111

CHAPTER III, RELAXATION OF OLIGONUCLEOTIDES STUDIED BY NMR..... 113

1. INTRODUCTION..... 113

2. EXPERIMENTAL METHODS..... 115

3. RESULTS..... 130

A) Saturation Recovery Experiments..... 130

B) Lifetimes of the Imino Protons for dCA_5G + dCT_5G at 5°C..... 136

C) Temperature Dependence of the Relaxation Rates for dCA_5G + dCT_5G..... 136

D) Lifetimes of the Imino Protons for the Hybrid Helix at 5°C..... 141

4. DISCUSSION..... 141

A) Are These Helices in the Open Limited Region?.. 145

5. CONCLUSIONS..... 150

PART II, OPTICAL STUDIES OF ETHIDIUM BINDING TO OLIGONUCLEOTIDES..... 152

CHAPTER IV, SPECTROSCOPIC AND THERMODYNAMIC STUDIES OF ETHIDIUM BINDING TO OLIGONUCLEOTIDES..... 152

1. INTRODUCTION..... 152

2. OPTICAL STUDIES EXPERIMENTAL..... 158

A) Materials..... 158

B) Methods..... 158

3. RESULTS

A) Determination of Equilibrium Constants of Ethidium Binding to Oligonucleotides..... 160

B) Stoichiometry of the Complex..... 171

C) Equilibrium Constants for Ethidium Binding to Oligonucleotides..... 172

D) Binding of Ethidium to Single Strands..... 174

E) Determination of the Enthalpy for Ethidium Binding to Oligonucleotides..... 175

F) Equilibrium Sedimentation Studies..... 176

G) Induced CD of Ethidium upon Binding to Oligonucleotides..... 179

H) Induced CD of Ethidium upon Binding to E. coli DNA..... 188

4. DISCUSSION AND CONCLUSIONS..... 191

A) Thermodynamics of Ethidium Binding to Oligonucleotides..... 191

B) Induced Circular Dichroism of Ethidium..... 193

APPENDIX I, NMR STUDY OF A DOUBLE HELIX WITH A MISMATCHED BASE..... 196

APPENDIX II, COMPUTER PROGRAMS..... 200

REFERENCES..... 203

CHAPTER I

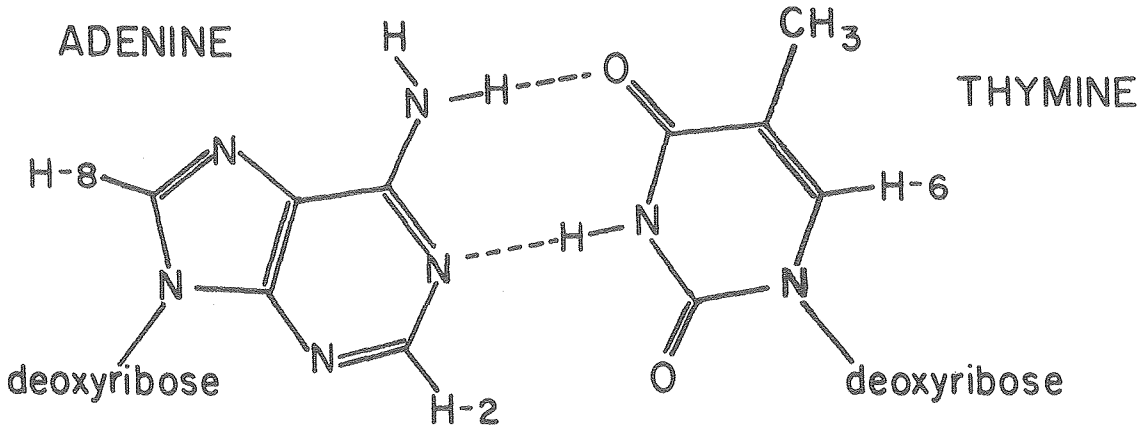
INTRODUCTION

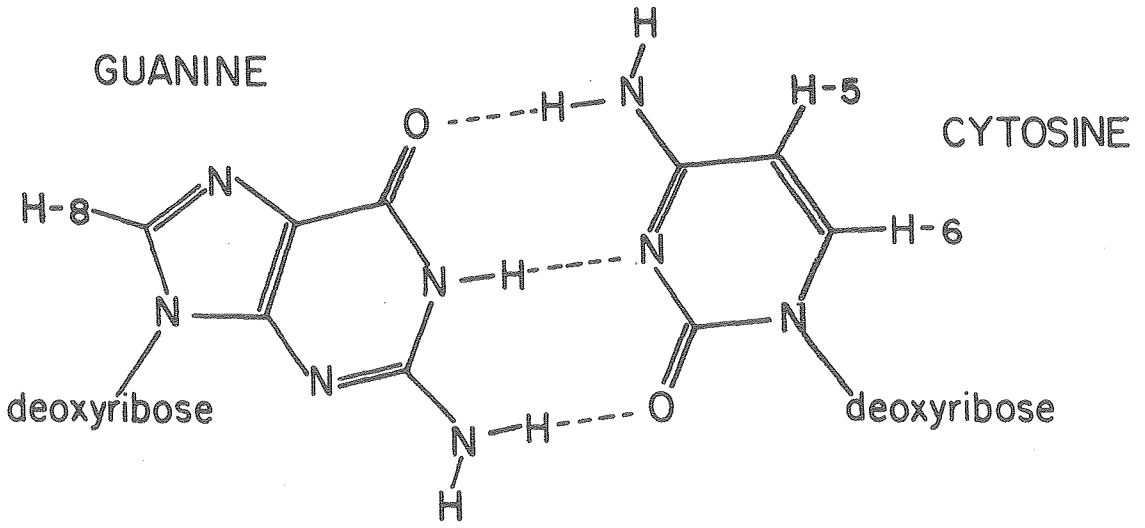
DNA carries all the genetic information in an organism. The structure of DNA was first proposed by Watson and Crick from X-ray diffraction studies.[1] Later X-ray fiber studies on RNA and hybrid polynucleotides showed there were large differences in the conformations of various types of nucleic acids.[2,3] In order to understand their structure and function in solution, nucleic acids have been studied by many techniques including absorbance, CD, NMR, and fluorescence.[4] Most of the early work concentrated on polymers, but as more detailed questions about structure and function were asked, shorter oligonucleotide systems were studied.

In this work we have used model oligonucleotides to study the properties of nucleic acids in solution. The oligonucleotides include a DNA duplex, an RNA duplex, a DNA-RNA hybrid duplex, and a DNA-RNA hybrid triplex, all of the same sequence. This gives us an opportunity to study the differences in these nucleic acid helices without having to worry about sequence effects. The main tool used was proton nuclear magnetic resonance. Figure 1.1 shows an A·T and C·G base pair, and the numbering of the base protons. Figure 1.2 shows a proton NMR spectrum for a typical oligonucleotide, and indicates the chemical shifts of protons

Figure 1.1. Watson-Crick base pairs in DNA.

WATSON - CRICK BASE PAIRS

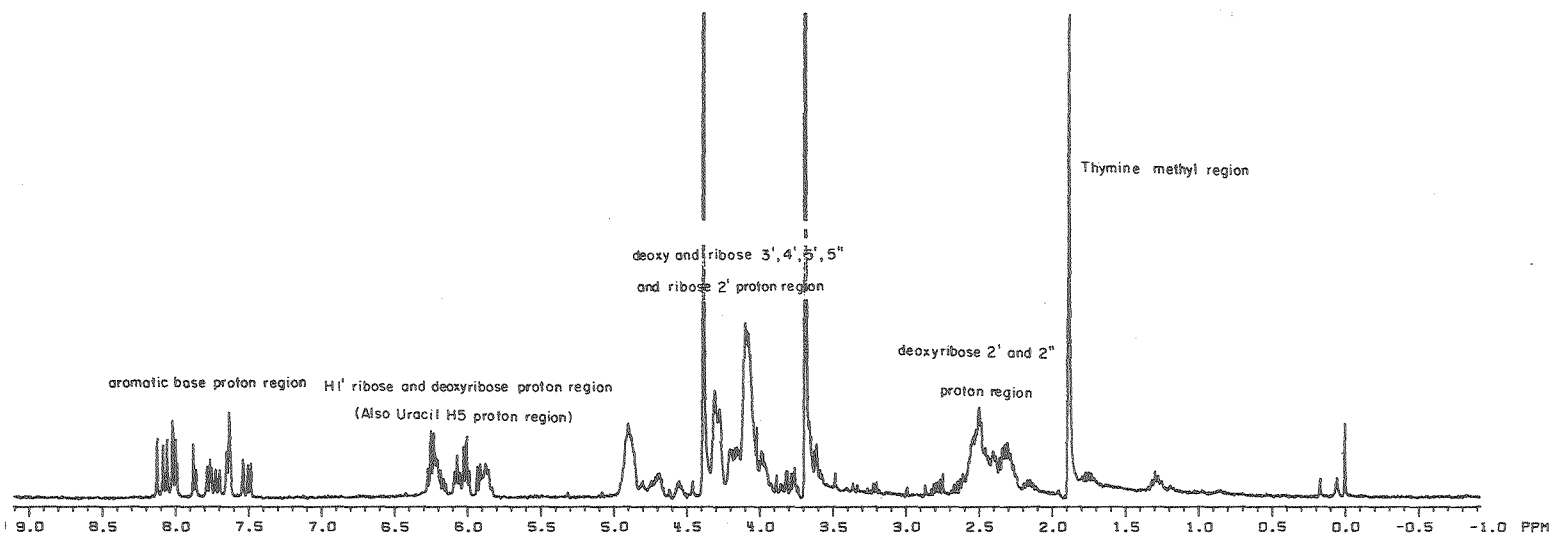




XBL 801-7722

Figure 1.2. ^{1}H NMR spectrum of $dCA_5G + dCT_5G$.

360 MHz Proton Spectrum of $dCA_5G + dCT_5G$ at 65°C



XBL 805-9521

on different parts of the nucleic acid. The chemical shifts and coupling constants of these protons were used to obtain information on the oligomers.

The aromatic bases in nucleic acides give rise to ring currents which affect the chemical shifts of nearby protons.[5] The calculated shapes and magnitudes of these ring currents for the four normal bases in DNA and RNA are shown in Figures 1.3a and 1.3b (thymine and uracil are assumed to have the same ring current in this work). The chemical shift of a particular proton will change with different conformations of the neighboring bases. Thus structural information can be obtained from the chemical shifts of protons in different parts of the helix. This method has been used in the past few years to obtain a much better understanding of the solution structures of tRNA's.[6,7]

The coupling constants of protons on the ribose or deoxyribose sugars of nucleic acids have been used by many workers to investigate the structures of small oligonucleotides.[8,9] The coupling constants are related to the dihedral angle between two protons by the Karplus equation.[10] We have used this empirical relation to obtain information on the sugar pucker of several oligonucleotides.

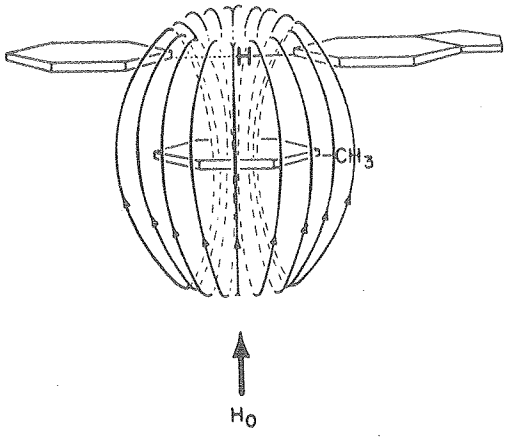
The second part of this work involves an investigation of the binding of ethidium to nucleic acids. Ethidium bromide is an aromatic dye which is known to cause frameshift

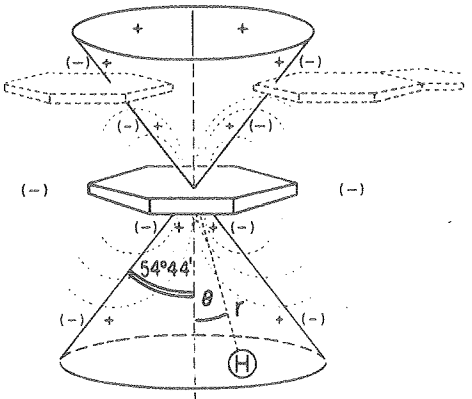
Figure 1.3.

a) Shape of ring currents due to bases in nucleic acids.

b) Magnitude of ring currents due to four bases.

a)



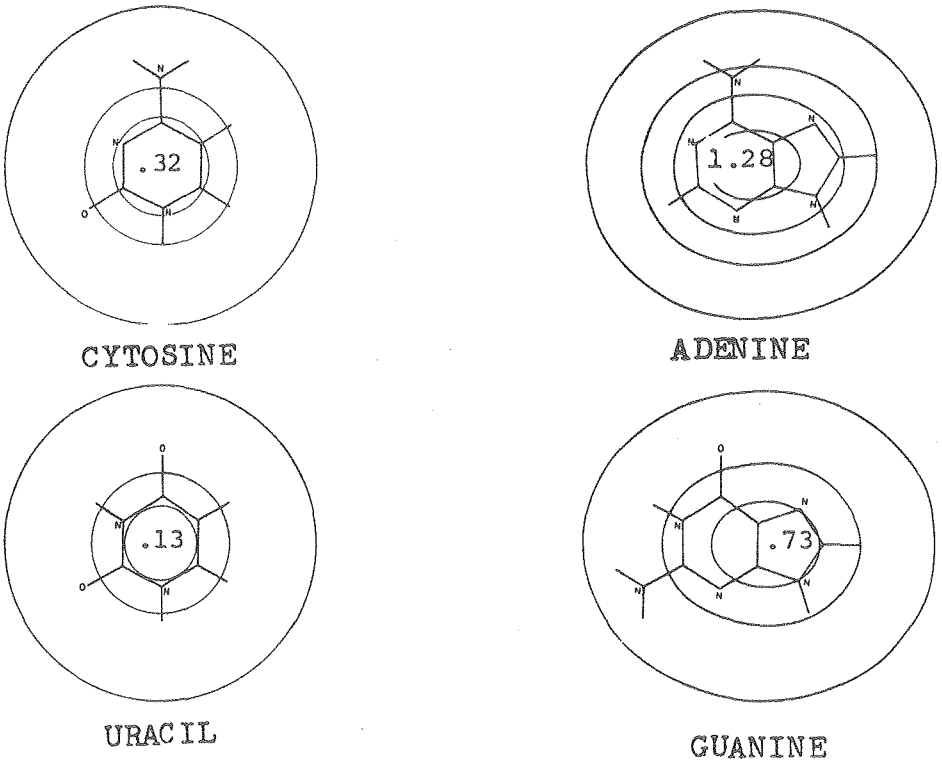


$$\frac{3\cos^2\theta - 1}{r^3}$$

XBL 805-9519

b)

SHIELDING VALUES DUE TO RING CURRENT OF BASE



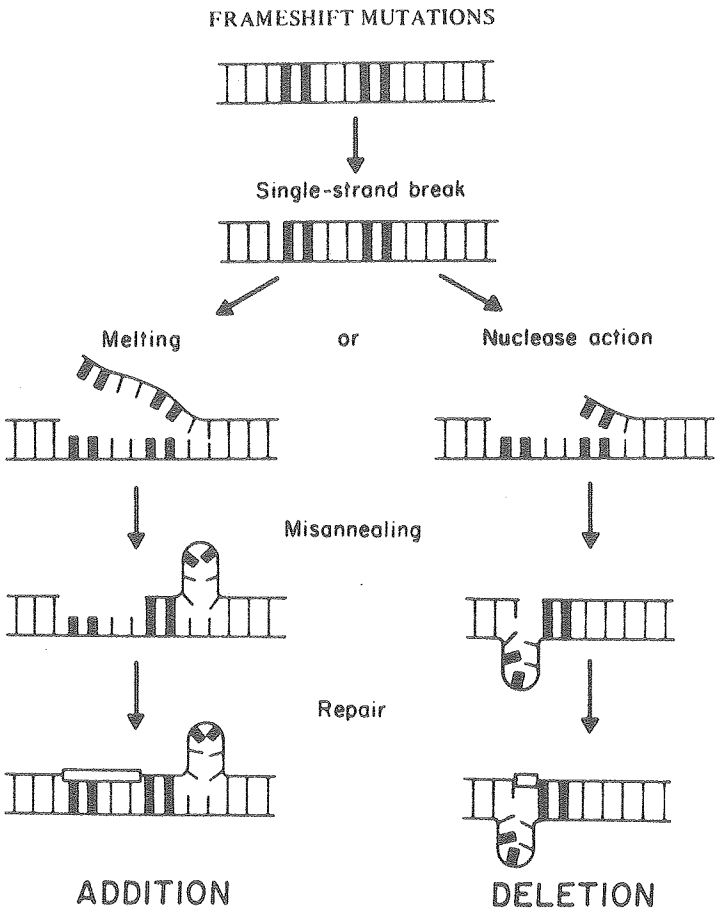
Numbers are for a plane 3.4 Å above or below the base.

XBL 805-9527

mutations in DNA.[11] By studying the physical properties of ethidium binding, we may be able to understand its mutagenic action. A model for frameshift mutagenesis has been proposed by Streisinger where the stabilization of a bulge in one strand causes the addition or deletion of a few base pairs during repair or replication of the DNA.[12] This mechanism is schematically shown in Figure 1.4; chemical mutagens are thought to stabilize these structures thus making frameshift mutations more likely.

Ethidium binding to three types of oligonucleotides was studied a) perfect helices (rCpG, or dCpG), b) complexes with a bulge on one strand (rGpUpG + rCpC), and c) complexes with bulged bases on both strands (rCpUpG). These represent possible structures involved in the Streisinger frameshift mechanism. The thermodynamic and spectroscopic properties of these complexes were compared.

Figure 1.4 A representation of the Streisinger model of frameshift mutagenesis.



XBL 782-7420A

CHAPTER II

1. INTRODUCTION

Proton nuclear magnetic resonance studies of oligonucleotides have greatly increased our understanding of the conformations and physical properties of nucleic acids in solution. The chemical shifts of the NH-N imino resonances, base proton resonances, and the H1' ribose or deoxyribose resonances have been used separately or in combination, to deduce RNA A or A'[3] type conformations for double helical ribo-oligonucleotides.[13-16] Double stranded deoxyribo-oligonucleotides have shown spectra consistent with a B form geometry in solution.[17-20] These conclusions are based mainly upon comparison of the experimental chemical shifts with those computed from ring current effects for an assumed geometry, as well as the sugar pucker deduced from the H1' proton coupling constants.

X-ray studies of RNA-DNA hybrid duplexes have shown them to adopt A type geometries.[3,21] The only ^{1}H NMR study of an RNA-DNA duplex shows the structure of the hybrid to be different than the DNA-DNA duplex of the same sequence, and consistent with an A form in solution.[22]

Triple stranded structures are not uncommon in polynucleotide or oligonucleotide solutions.[4] Under conditions where triplexes were formed, Geerdes and Hilbers observed

the low field imino proton region of [oligo A-(oligo U)$_2$].[2,3] They saw separate resonances for the Watson-Crick and reverse Hoogsteen base pairs. Kallenbach et al. observed similar results for the other triplexes including [AMP-(oligo U$_{15}$)$_2$].[19]

In this work we report the ^{1}H NMR of oligonucleotides in a DNA duplex, an RNA duplex, a DNA-RNA hybrid duplex, and a DNA-RNA hybrid triplex, of the same sequence. The molecules are (I) dCT$_5$G + dCA$_5$G, (II) rCU$_5$G + rCA$_5$G, (III) dCT$_5$G + rCA$_5$G, and (IV) rCU$_5$G + dCA$_5$G. The first three molecules form duplexes which all have different conformations in solution; IV forms a triple strand which is much less stable than the other structures. These results are consistent with optical studies previously reported on these molecules, which pointed out the importance of the instability of the DNA-RNA hybrid present at termination of transcription.[24]

The thermodynamics of oligonucleotides have been studied by optical methods as well as by proton NMR.[25] The melting temperatures, T_m's, of the helix-to-coil transition from both methods have agreed fairly well.[15,19,26] Melting of the end base pairs before the rest of the helix has been seen in several of the systems with A·U or A·T base pairs on the ends of the helix.[15,19,27] In helices with C·G base pairs on the ends there seems to be no[16] or little[27] differential melting of the ends of the helix before the rest of the oligonucleotide. However, most of the systems studied thus

far have been self-complementary oligonucleotides (except Hughes et al.[16]); this precludes the measurement of the temperature dependence of the single strands. Ignorance of the temperature dependence of the single strands throughout the melting transition can lead to errors in the apparent T_m's; these errors will usually be larger for the interior base pairs than for the terminal ones. The oligonucleotides studied here are not self-complementary, so the single strand dependences have been observed. The measured T_m for each base proton is thus a more accurate representation of the melting of that part of the helix.

2. MATERIALS AND METHODS.

A) Oligonucleotides

The deoxyribo-oligonucleotides were synthesized by the diester method of Khorana.[28] The ribo-oligonucleotides were enzymatically prepared with polynucleotide phosphorylase.[25] The oligonucleotides were synthesized and kindly provided by Dr. Frank Martin and Mr. David Koh. Separation and purification of the oligomers were performed by RPC-5 column chromatography. Desalting of the samples was performed on Biogel P-2 columns (Bio-Rad). All samples were run in 8.0 mM Na_2HPO_4, 20 mM NaH_2PO_4, 0.18 M NaCl, 0.1 mM Na_2EDTA, pH = 7.0, unless otherwise noted. Concentrations of the oligonucleotides were calculated from the absorbance of the solutions at 260 nm. The extinction coefficients were calculated from extinction coefficients of dinucleoside monophosphates and mononucleotides

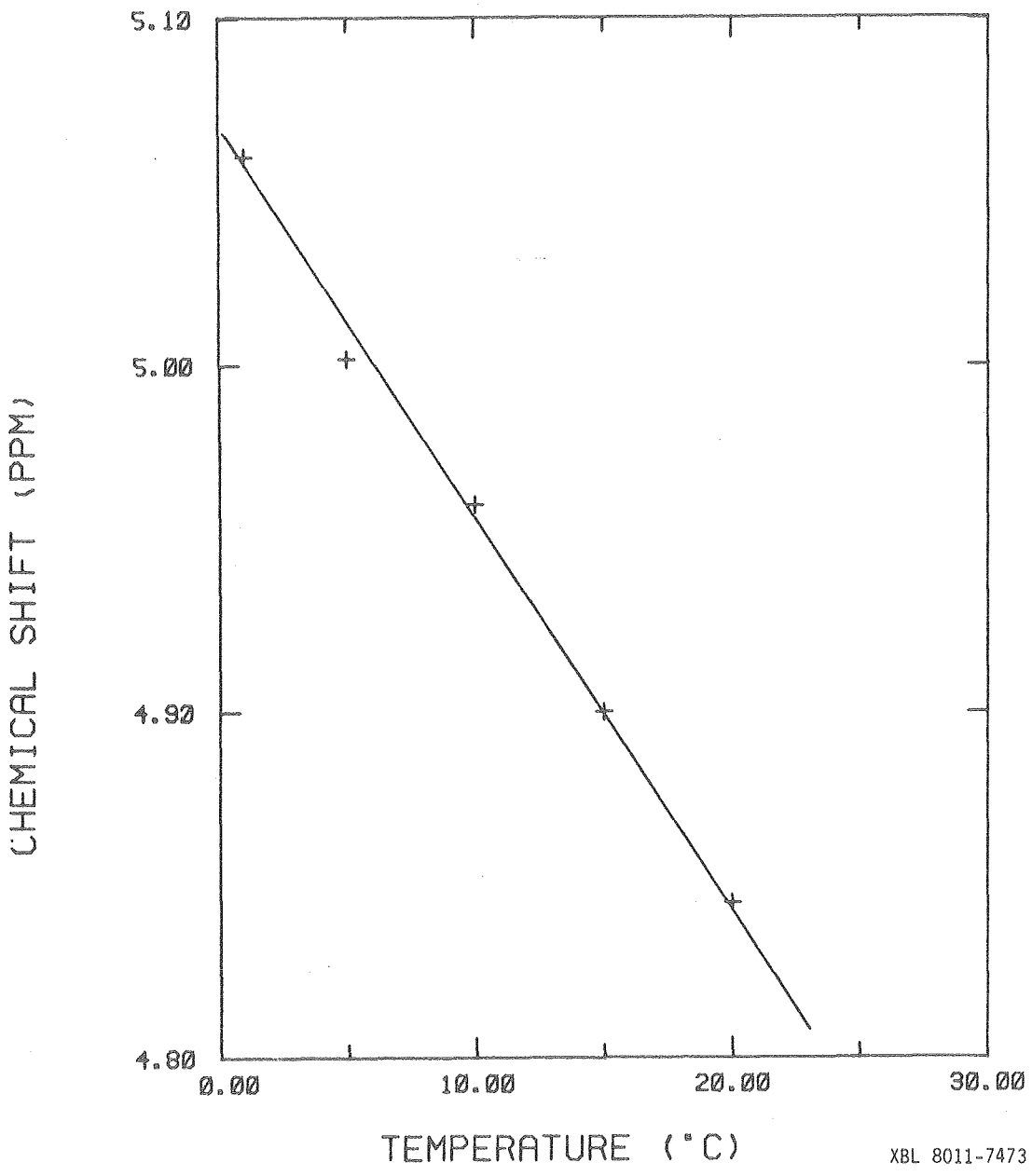
with the assumption of only nearest neighbor interaction.[29] The values obtained in this way are 79 x 10^3, 58 x 10^3, 79 x 10^3, and 66 x 10^3 for dCA_5G, dCT_5G, rCA_5G and rCU_5G, respectively, at 25°C.

B) Experimental NMR

NMR spectra were taken on the HXS-360 MHz instrument at Stanford Magnetic Resonance Laboratory. Temperature was controlled to ± 1°C by a B-ST 100/700 Bruker temperature controller. Spectra of the non-exchangeable protons were measured in D_2O in the normal Fourier transform mode using 5 mm NMR tubes (Wilmad). Spectra of the exchangeable protons were measured in H_2O using the Redfield 214 pulse sequence to minimize the water signal.[30] For these samples 5 mm microtubes which hold 160 µl of solution (508 CP Wilmad) were used. All solutions were measured at concentrations of either 1.0 mM or 0.5 mM per strand. Spectra measured in D_2O were all referenced to the internal standard TSP, while the spectra in H_2O were referenced to the H_2O peak. The chemical shifts obtained in this way are accurate to ± 0.005 ppm for the D_2O and ± 0.05 ppm for the H_2O studies.

The temperature dependence of the chemical shift of H_2O relative to TSP was calibrated in our buffer. The sample used for calibration was our standard buffer in pure H_2O plus 20 mM TSP as reference. A normal Fourier transform spectrum was taken with a short pulse length and

Figure 2.1. Chemical shift of H_2O with temperature for 8 mM Na_2HPO_4, 20 mM NaH_2PO_4, 0.18 M NaCl, 0.1 mM EDTA, pH = 7.0.


5.10
5.00
4.90
4.80
0.00
10.00
20.00
30.00
CHEMICAL SHIFT (PPM)
TEMPERATURE (°C)
XBL 8011-7473

low power. The chemical shift of H_2O is linear with temperature over this temperature range, and is shown in Figure 2.1. The chemical shift of H_2O was used as a reference in the Redfield experiments where it was usually impossible to see the TSP peak. Since there is a large distortion around the H_2O peak in the Redfield pulse sequence, it was difficult to get accurate values for the chemical shift of H_2O in the spectrum. This problem was eliminated by also taking a short pulse, low power spectrum of the sample under conditions identical to those used in the Redfield experiment. In the short pulse experiment there are no distortions of the H_2O peak, so its position in the Redfield spectrum was taken as the same as that observed in the short pulse experiment. This method was tested and found to obtain the chemical shift of the H_2O peak to better than ± 0.05 ppm in the Redfield pulse sequence spectra.

All data were collected on a Nicolet 1180 computer with 16K data points and a sweep width of ± 1800 Hz for the D_2O work, and 8K data points and a sweep width of ± 5000 Hz for the H_2O spectra. Spectra were taken every 10° from 35 to 65°C and every 5° from 5 to 35°C, for the spectra in D_2O. In H_2O measurements were made every 5° from 5 to 25°C or until peaks were no longer observed.

Samples in D_2O were prepared by lyophilizing three times in 99.8% D_2O (Bio-Rad) and then dissolved in 100% D_2O (Bio-Rad). The internal reference TSP was added to the sample after the addition of the 100% D_2O, since it had a tendency to remain insoluble after lyophilization with the oligonucleotides. Samples of the component oligomers titrated with Mn^{2+} were measured in D_2O with no buffer added at pD = 7.0. The pD was calculated by adding 0.4 to the measured pH, with the pD adjusted by the addition of aqueous NaOD or DCl.

3. RESULTS

A) Assignments of Nonexchangeable Protons

The assignments of the chemical shifts of the base protons were accomplished on the single strands. The method of incremental assignment was used at 65°C where there is less stacking of the bases in the single strands.[15] The chemical shifts at other temperatures were obtained from the change in peak position with temperature. Chemical shifts of the single strand oligomers at 65°C are given in Table I.

The aromatic region (6.6-8.5 ppm) of rCU_5G was poorly resolved due to the H6 doublets on the uracil residues, so no assignments were made in solutions which contained this oligomer. In the other solutions, the cytosine resonances were the only doublets in the aromatic region and easily identified. The adenine H8 protons were differentiated from the adenine H2 protons because of the longer T_1 of the H2

TABLE I

CHEMICAL SHIFTS OF SINGLE STRANDS AT 65°C*

dCA_5G	C#1	A#2	A#3	A#4	A#5	A#6	G#7
AH2	--	7.985	(7.752	7.725	7.696)	7.854	--
CH6	7.489	--	--	--	--	--	--
A or GH8	--	8.120	(8.081	8.052	8.016	8.012)	7.872

rCA_5G	C#1	A#2	A#3	A#4	A#5	A#6	G#7
AH2	--	8.091	(7.982	7.971	7.953)	8.034	--
CH6	7.632	--	--	--	--	--	--
A or GH8	--	8.265	(8.175	8.165	8.152)	8.190	7.878

dCT_5G	C#7	T#6	T#5	T#4	T#3	T#2	G#1
C or TH6	7.770	7.532	(7.627	7.627	7.627)	7.645	--
CCH_3	--	(1.878	1.878	1.878	1.878	1.878)	--
GH8	--	--	--	--	--	--	7.993

*Values in parenthesis indicate that we were unable to assign these peaks to particular bases in the sequence.

protons.[31] The H8 on guanine was distinguished from H8 of adenine by the fact that it exchanges much faster upon heating in D_2O at 80°C. (In one hour at 80°C the guanine H8 intensity was decreased over 50% while the adenine H2 was decreased by about 10%.)

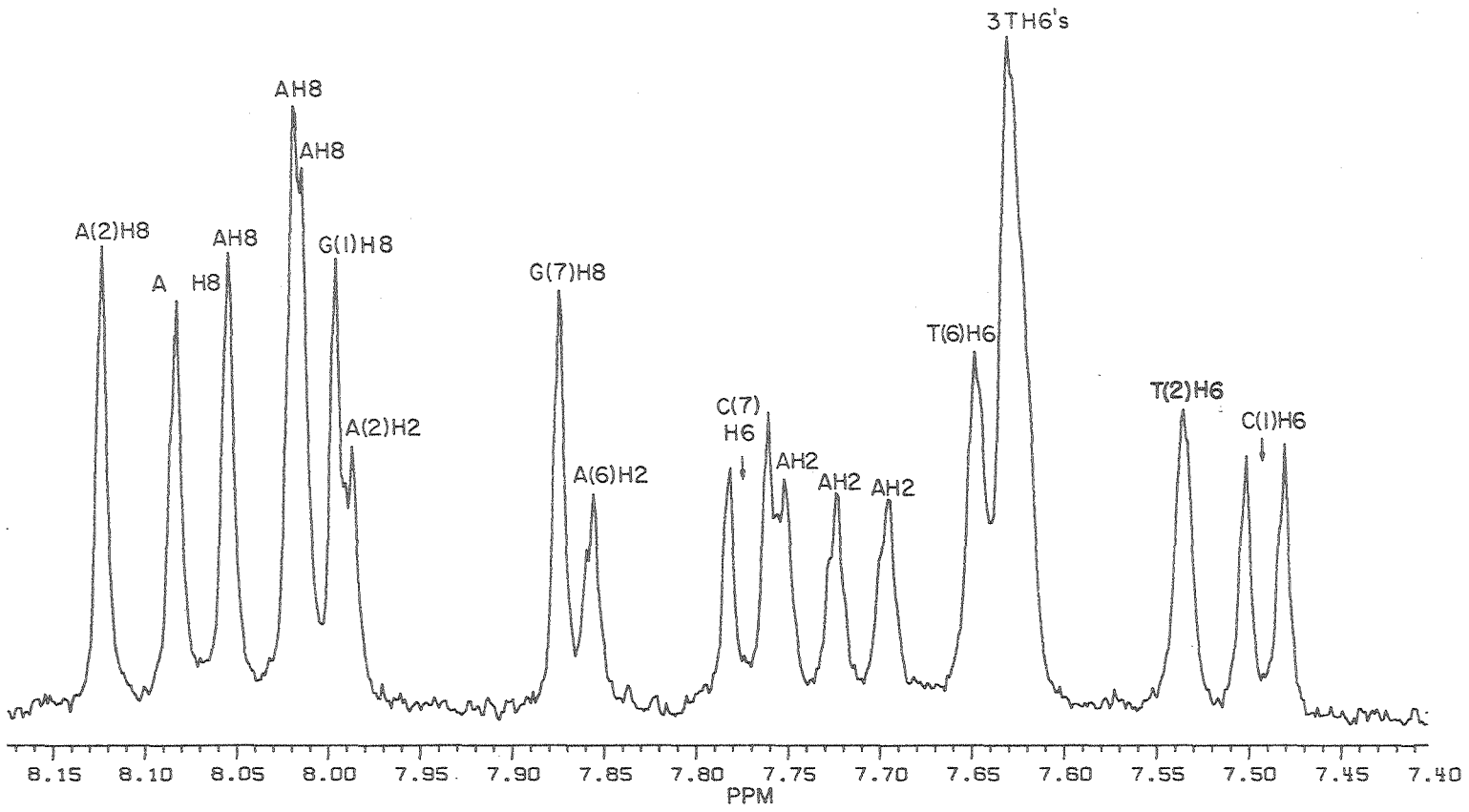
In the oligonucleotide dCA_5G, we unambiguously assigned all the base protons on the cytosine, the guanine, and the #2 adenine. Figure 2.2 shows the base proton spectrum of dCA_5G + dCT_5G in the single strands at 65°C, and also defines the numbering scheme. For the other penultimate adenine (#6), we were able to specifically assign the H2, but not the H8 proton. The four H8 adenine resonances which belong to the three internal adenines (#3-5) and the #6 adenine were identified, but not assigned to specific bases. The H2 protons on the three internal adenines all have very similar chemical shifts and we were unable to assign resonances to particular adenines in the sequence.

In order to make assignments of the penultimate adenine resonances (#2 and #6) it was necessary to study the smaller components of dCA_5G. The following compounds at 65°C were used for comparison with dCA_5G: dCpA, dCpApA, $dC(pA)_6$, dpApA, $d(pA)_4$, dpApG, and dpApApG. The assignments of the dCpA and dpApG were made by inspection. Tentative assignment of the base protons on dApA has previously been made by Chang and Sarma.[32] The assignment of the base protons in

Figure 2.2. Assignments of base protons in single strands of dCA_5G + dCT_5G at 65°C.

360 MHz ^{1}H NMR of
$dCA_5G + dCT_5G$, 1 mM/strand, 65°C
D_2O, 0.22 M Na^+, pH = 7.0

SINGLE STRANDS
5'CAAAAAG3' + 5'CTTTTTG3'
1234567 7654321



XBL 801-7675

dpApA required titration with Mn^{2+}, which preferentially binds to the terminal phosphate at pD = 7 and thus broadens the 5' adenine resonances relative to those of the 3' adenine.[33] The broadening of the NMR signal from the paramagnetic Mn^{2+} ion is due to the electron-spin, nuclear-spin, dipole-dipole interaction.[34,35] This interaction goes as $1/r^6$ so only protons located very close to the phosphate binding site are broadened.[36]

The linewidths of the dpApA base protons during the titration are given in Table II. The H8 proton at 8.33 ppm clearly broadens before the H8 proton at 8.28 ppm thereby assuring the assignment of the 8.33 ppm resonance to the 5' residue, and the 8.28 ppm peak to the 3' residue. The H2 resonances were more difficult to assign because both adenine H2 protons are fairly far from the 5' phosphate. (See the single stranded right handed stack, shown in Figure 2.3, for an estimate of the distance between the 5' phosphate and the two adenine H2 protons.) At 50 μM Mn^{2+} the H2 proton at 8.03 ppm broadened to 12 Hz whereas the H2 resonance at 8.10 ppm only broadened to 6 Hz. This is indicative that the 8.03 peak belongs to the 5' adenine.

The assignment of the dpApA helped in the assignment of the oligomers dpApA and dCpApA. Only partial assignment of the oligomers $dC(pA)_6$, and $d(pA)_4$ were possible. The chemical shifts of all the components at 65°C are given in Tables III - V.

TABLE II

LINEWIDTHS OF THE BASE PROTONS OF dpApA WHEN TITRATED WITH Mn^{2+} AT pD = 7.0

	H8 protons		H2 protons	
Chemical shift (ppm)	8.333	8.281	8.094	8.029
$[Mn^{2+}]$				
0.0	3.0	3.0	3.0	3.0
3.0 μM	12.0	6.0	3.0	3.8
10 μM	24	15	2.2	3.0
50 μM	a	a	6	12

[a]Linewidth too broad to measure.

TABLE III

CHEMICAL SHIFTS OF THE COMPONENTS OF dCA_5G AT 65°C[a]

1 2
dCpA

C(1)		A(2)	
H6	H5	H8	H2
7.58	5.98	8.38	8.24

1 2 3
dCpApA

C(1)		A(2)		A(3)	
H6	H5	H8	H2	H8	H2
7.52	5.94	8.31	8.10	8.20	8.05

1 2 I I I I 7
dCpApApApApApA

C(1)	A(2)		A(I or 6)
H6	H8	H2	H8
7.5	8.12*	7.97*	(8.12,8.09,8.06,8.05,8.01)
			H2
			(7.97,7.81,7.78,7.73,7.73)

1 2
dpApA

A(1)		A(2)	
H8	H2	H8	H2
8.33	8.03*	8.24	8.10*

1 2 3 4
dpApApApA

A(1)		A(2 and 3)		A(4)	
H8	H2	H8	H2	H8	H2
8.25*	7.98	8.10	7.80	8.20*	7.98

1 2
dpApG

A(1)		G(2)
H8	H2	H8
8.20	8.20	7.99

1 2 3
dpApApG

A(1)		A(2)		G(3)
H8	H2	H8	H2	H8
8.22	8.07	8.18	7.97	7.90

TABLE III

(continued)

*Indicates only tentative assignments.

[a]Parentheses indicates that we were unable to make assignments to particular bases in the sequence.

TABLE IV

CHEMICAL SHIFTS OF THE COMPONENTS OF dCT_5G at 65°C

1 2
dCpT

C(1)		T(2)
H6	H5	H6
7.80	6.02	7.68

1 2 3
dCpTpT

C(1)		T(2 or 3)
H6	H5	H6
7.77	6.02	7.65

1 2 3 4 5
dCpTpTpTpT

C(1)		T(2)	T(3-5)
H6	H5	H6	H6
7.78	--	7.66	7.65

$dC(pT)_6$

C(1)		T(all)
H6	H5	H6
7.77	--	7.63

dpTpG

T(1)		G(2)
H6	H5	H8
7.59	--	8.01

1 2 3
dpTpTpG

T(1)	T(2)	G(3)
H6	H6	H8
7.69	7.55	8.00

TABLE V

COMPONENTS OF rCA_5G AT 65°C[a]

1 2			
rCpA			
C(1)		A(2)	
H6	H5	H8	H2
7.76	6.00	8.41	8.29

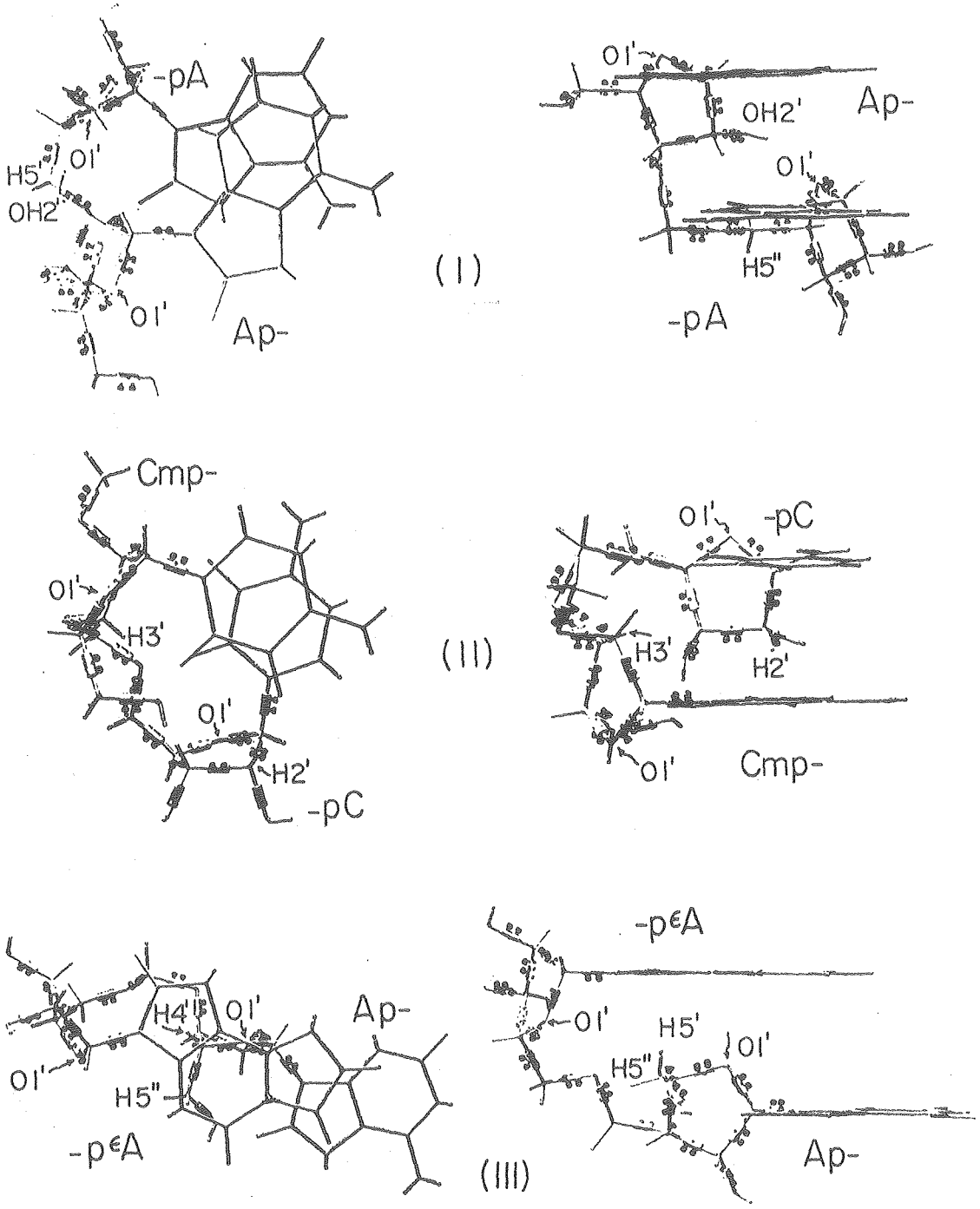
1 2 3					
rCpApA					
C(1)		A(2)		A(3)	
H6	H5	H8	H2	H8	H2
7.66	5.83	8.32	8.20*	8.32	8.14

1 2 3 4 5 6					
rCpApApApApA					
C(1)		A(2)		A(3-5)	
H6	H5	H8	H2	H8	H2
7.63	5.82	8.29	8.16	(8.23)	(8.07,8.06,8.03)

*Indicates only tentative assignments.

[a]Parentheses indicates that we were unable to make assignments to particular bases in the sequence.

Figure 2.3 Solution conformations of dinucleotides proposed by Lee and Tinoco.[37] Figure (I) shows the conformation for a right handed stack of rApA.

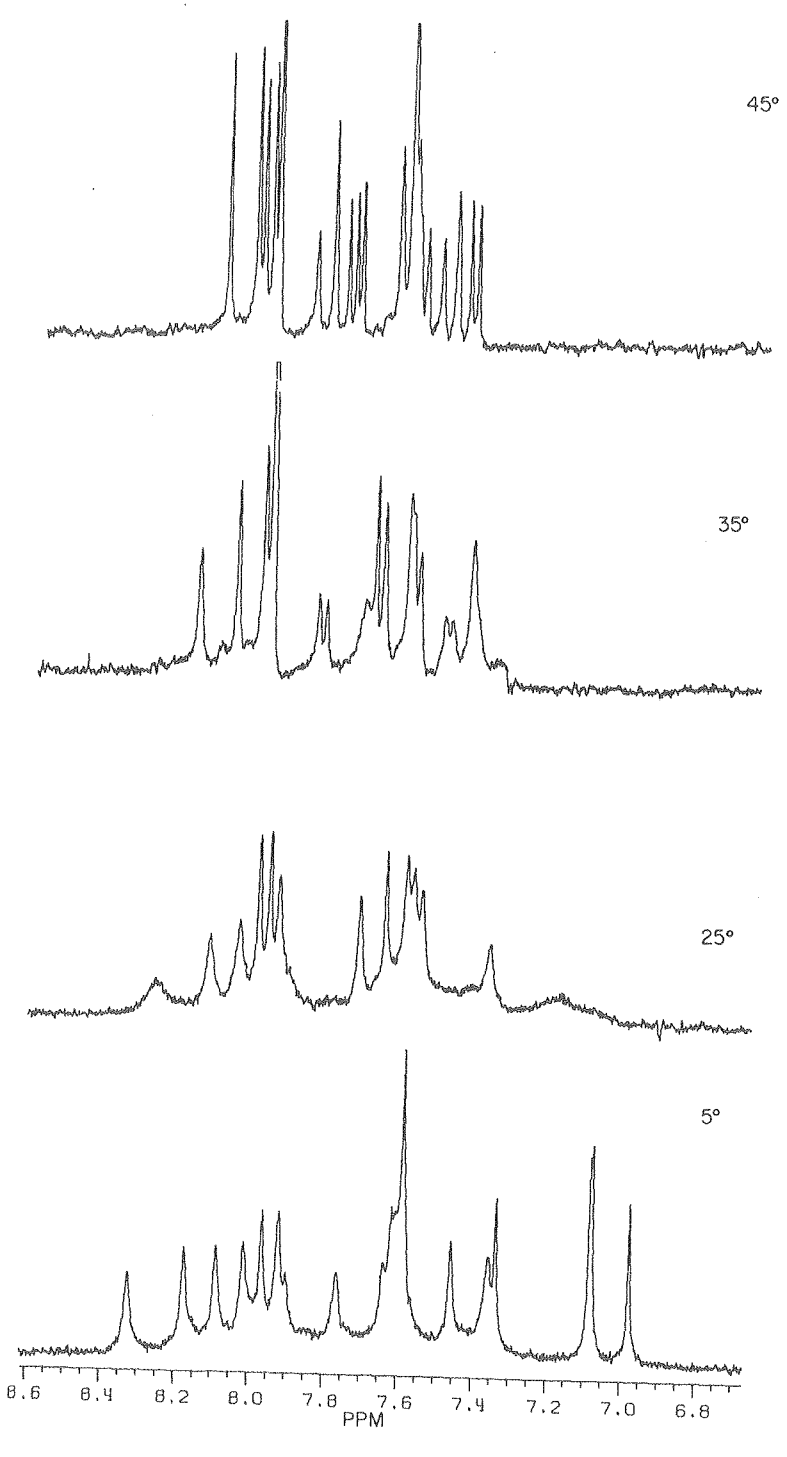

-pA
H5'
O1'
OH2'
O1'
Ap-
(I)
O1'
OH2'
Ap-
O1'
H5"
-pA
Cmp-
O1'
H3'
O1'
H2'
-pC
(II)
O1'
-pC
H3'
H2'
O1'
Cmp-
H4'
O1'
Ap-
O1'
H5"
-pεA
(III)
-pεA
O1'
H5'
O1'
H5"
Ap-

The assignments of the base protons in dCT_5G were made in a manner similar to that for dCA_5G. The components used were dCpT, dCpTpT, $dC(pT)_4$, $dC(pT)_6$, and dpTpTpG. Unambiguous assignments of the cytosine and guanine base protons as well as the H6 protons on the penultimate thymines were made. The three internal thymine H6 protons all had the same chemical shifts at 65°C. The assignments of the thymine methyl peaks to specific bases in the sequence were not made at any temperature. Chemical shifts of thes oligomers at 65°C are given in Table IV.

For the ribo-oligonucleotides no attempt was made to assign the uracil H5 and H6 base protons to particular uracils on the oligomers. Assignments of rCA_5G were made by comparison with the series $r(Ap)_nApG$ assigned by Shum,[51] as well as our analysis of the oligomers rCpA, rCpApA, and rCpApApApA (see Table V). From this work we were able to assign the cytosine and guanine base protons as well as the H8 and H2 on the penultimate adenines. Again, the assignments of the three internal H8 protons or the three internal H2 protons to specific bases were not attempted.

The assignment of the base protons at other temperatures was made by following each peak's shift with temperature; this is illustrated in Figures 2.4 and 2.5. Spectra were taken every ten degrees at higher temperatures where there are small changes in chemical shifts with temperature, and every 5°C in the range from 5 to 35°C, where the peaks shift

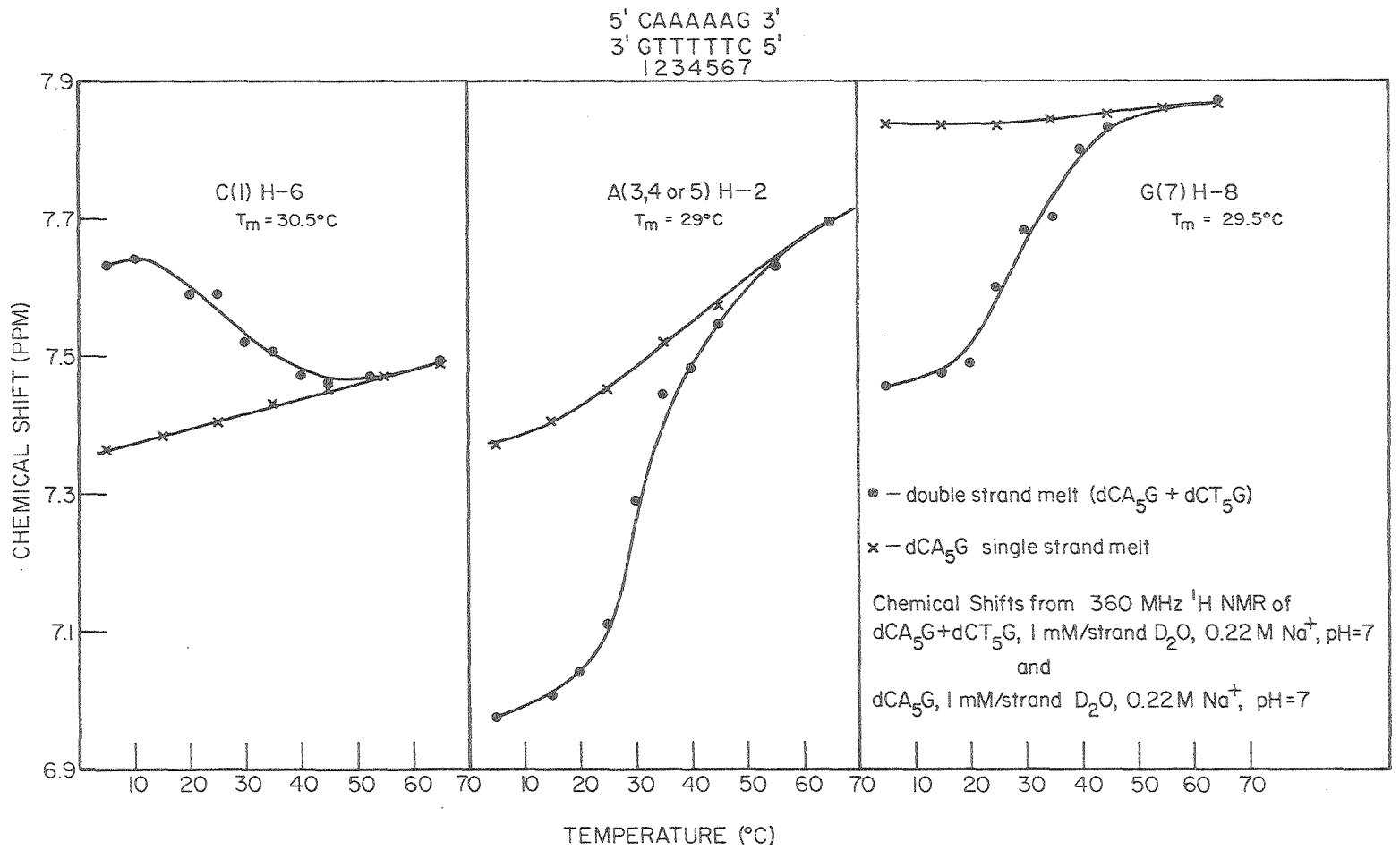
Figure 2.4. Helix-to-coil transition of dCA_5G + dCT_5G.


360 MHz ^{1}H NMR of
dCA_5G + dCT_5G, 1 mM/strand
D_2O, 0.22 M Na^+, pH = 7.0
45°
35°
25°
5°
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
PPM

XBL 801-7673

Figure 2.5. Melting curves for nonexchangeable base protons in $dCA_5G + dCT_5G$.

NMR MELTING CURVES: CHEMICAL SHIFT VS. TEMPERATURE



XBL 801-7671

dramatically with temperature and exchange broadening is evident. The assignment of the chemical shifts at 5°C for the double strand dCA_5G + dCT_5G was aided by comparison of spectra taken before and after the guanine H8 protons were exchanged in D_2O by heating at 75°C for one hour. Chemical shifts of this double strand are given in Table VI. A small four bond coupling of the thymine methyl protons with the H6 protons helps in the identification of the H6 resonances. Assignment of the adenine H2 protons was aided by the fact that these peaks were sharper than the rest of the resonances in the double strand, presumably because of their longer T_2 compared to the T_2 of the other protons. The base proton spectrum of dCT_5G + dCA_5G in the double strand at 5°C is shown in Figure 2.6.

The chemical shifts for the protons in the double strand hybrid helix (dCT_5G + rCA_5G) which could be unambiguously followed are given in Table VI. In the complexes rCA_5G + rCU_5G and rCU_5G + dCA_5G the spectra were so poorly resolved in the aromatic region because of the uracil doublets that we were unable to see individual peaks in the aromatic region at temperatures lower than 35°C, even at 360 MHz.

The H1' protons in deoxyribose and ribose sugars resonate at 5.5-6.5 ppm. The H1' proton in deoxyribose sugars is coupled to the H2' and H2" protons while the ribose H1' is only coupled to the H2' proton. Thus the H1' protons on

TABLE VI

EXPERIMENTAL AND CALCULATED CHEMICAL SHIFTS OF THE NONEXCHANGEABLE BASE PROTONS IN THE DOUBLE HELIX[a]

	#1 C G	#2 A T	#3 A T	#4 A T	#5 A T	#6 A T	#7 G C
$dCA_5G + dCT_5G$ helix							
AH2	--	7.337	(7.084	7.082	6.976)	7.590	--
C or T H6	7.630	7.356	(7.640	7.617	7.590)	7.760	7.910
G or A H8	7.964	8.324	(8.173	8.085	8.012	7.920)	7.870
$rCA_5G + dCT_5G$ helix							
AH2	--	7.766	(7.229	7.229	6.876)	7.317	--
C or T H6	b	7.860	(7.676	7.621	7.574)	7.502	∿8.0
G or A H8	b	8.163	b	b	b	b	b
Calculation for DNA B form							
AH2	--	6.78	6.92	6.99	7.07	7.73	--
C or T H6	7.93	7.83	7.84	7.84	7.85	7.83	7.94
G or A H8	8.14	8.40	8.32	8.29	8.30	8.30	7.96
Calculation for RNA A form							
AH2	--	6.89	7.12	7.15	7.17	7.61	--
C or T H6	8.13	7.72	7.73	7.72	7.77	7.77	8.10
G or A H8	8.05	8.46	7.89	7.83	7.83	7.80	7.98
Calculation for RNA A' form							
AH2	--	6.80	7.06	6.82	6.89	7.57	--
C or T H6	8.13	7.71	7.72	7.71	7.76	7.76	8.10
G or A H8	8.04	8.42	7.79	7.72	7.72	7.69	7.45

TABLE VI

(continued)

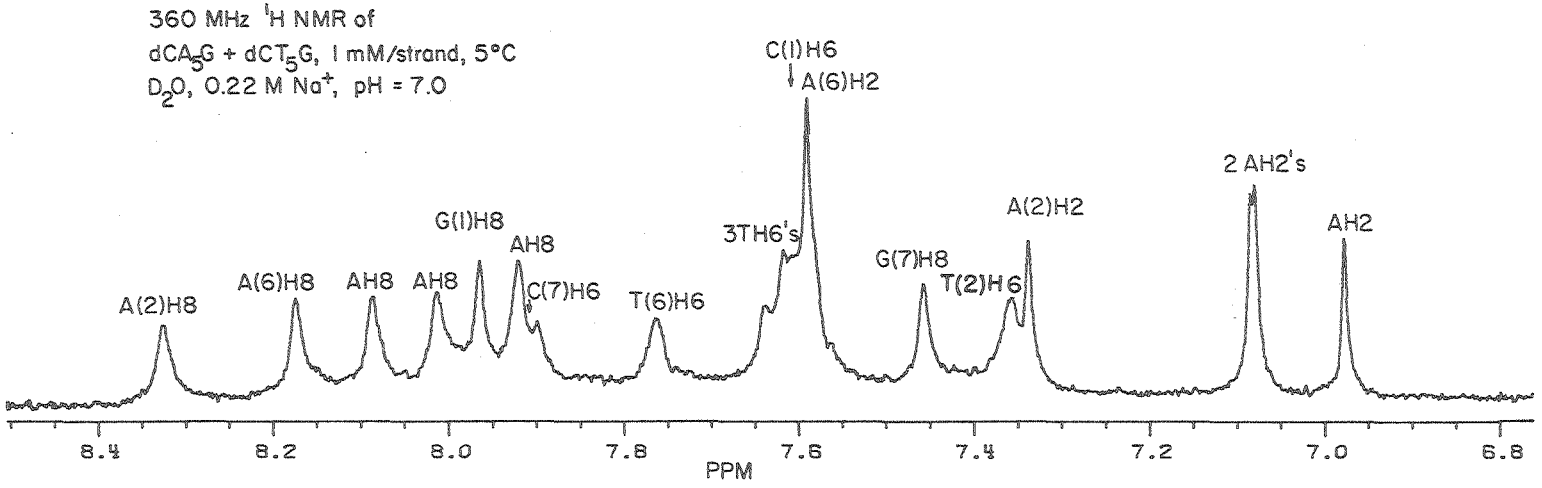
[a]The calculations were done using numbers dervied by Arter and Schmidt,[41] and we assumed thymine has the same ring current as uracil. Values in parenthesis indicate that we were unable to assign these peaks to particular bases in the sequence.

[b]Poor resolution made assignment of this proton impossible.

Figure 2.6. Assignments of nonexchangeable base protons in double strand of $dCA_5G + dCT_5G$.

DOUBLE STRAND

5'CAAAAAG3'
3'GTTTTTC5'
1234567



XBL 801-7674

the two different rings are easily differentiated. We were not able to assign the H1' protons to particular sugars in the sequence, but only to either deoxyribose or ribose sugars.

B) Assignment of the Base Paired Imino Protons

The base paired imino protons for the duplex dCT_5G + dCA_5G are shown in Figure 2.7a. Partial assignment of these resonances was made by the temperature dependences of the protons, as well as comparison with calculations of chemical shifts for the double strand. The guanine H1 proton involved in a hypothetical isolated base pair has been found to resonate 1 ppm upfield from the uracil or thymine H3 in the isolated base pair.[6,38] The numbers for these isolated base pairs derived from tRNA data are 14.4 ppm for the A·U base pair and 13.6 ppm for the C·G base pair.[6] Since there is only a limited amount of work done on the imino protons in deoxyribo-oligonucleotides of known sequence,[39] the number for isolated base pairs derived from the tRNA work have been applied to the DNA helices.

For the dCA_5G + dCT_5G helix there are clearly seven resonances in the imino region, as seen from the integration shown in Figure 2.8. The two resonances found at 12.93 and 13.55 ppm broaden before the other peaks. Figure 2.9 shows that at 20°C these two peaks are extremely broad

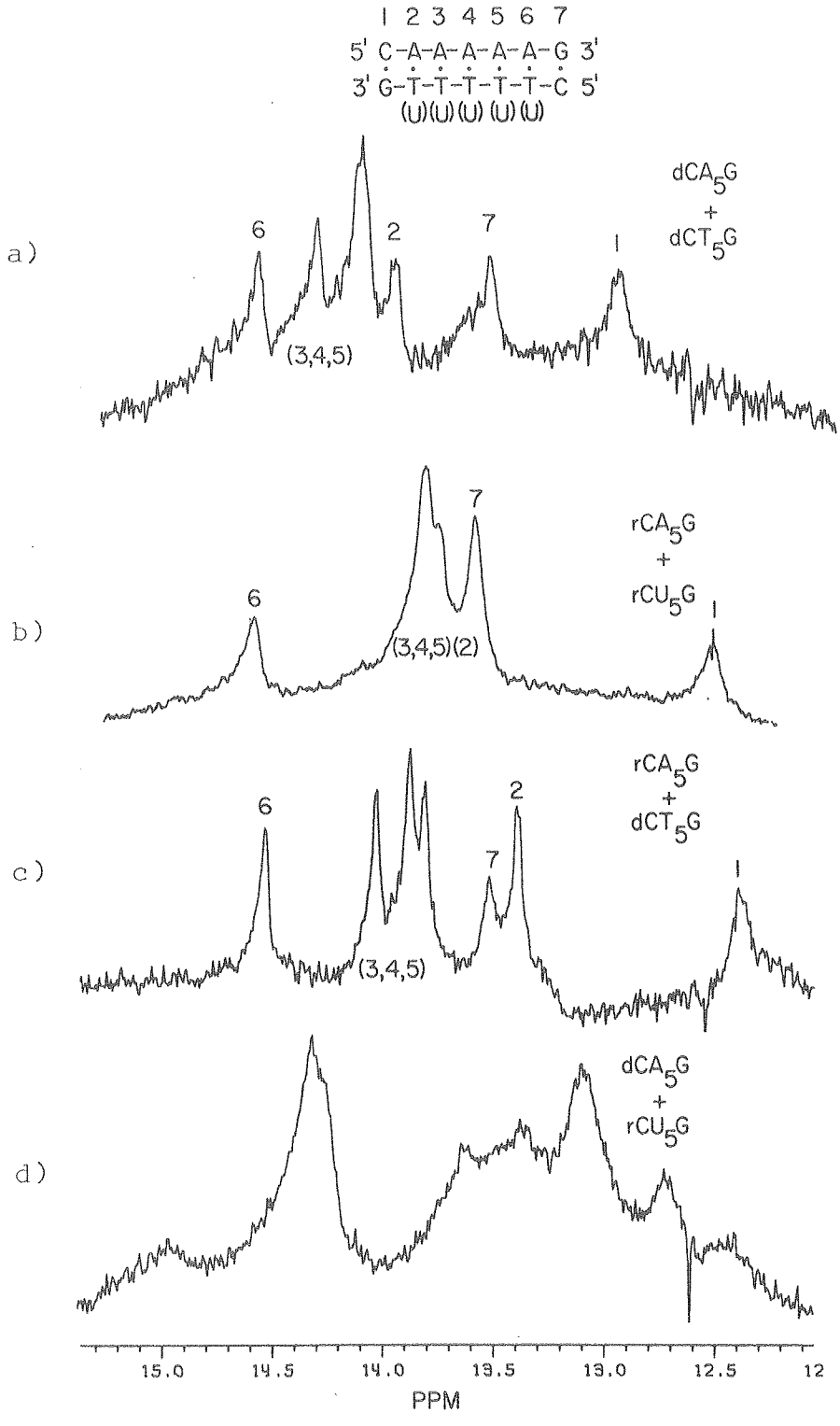
Figure 2.7.

a) dCA_5G + dCT_5G double strand at 5°C.

b) rCA_5G + rCU_5G double strand at 5°C.

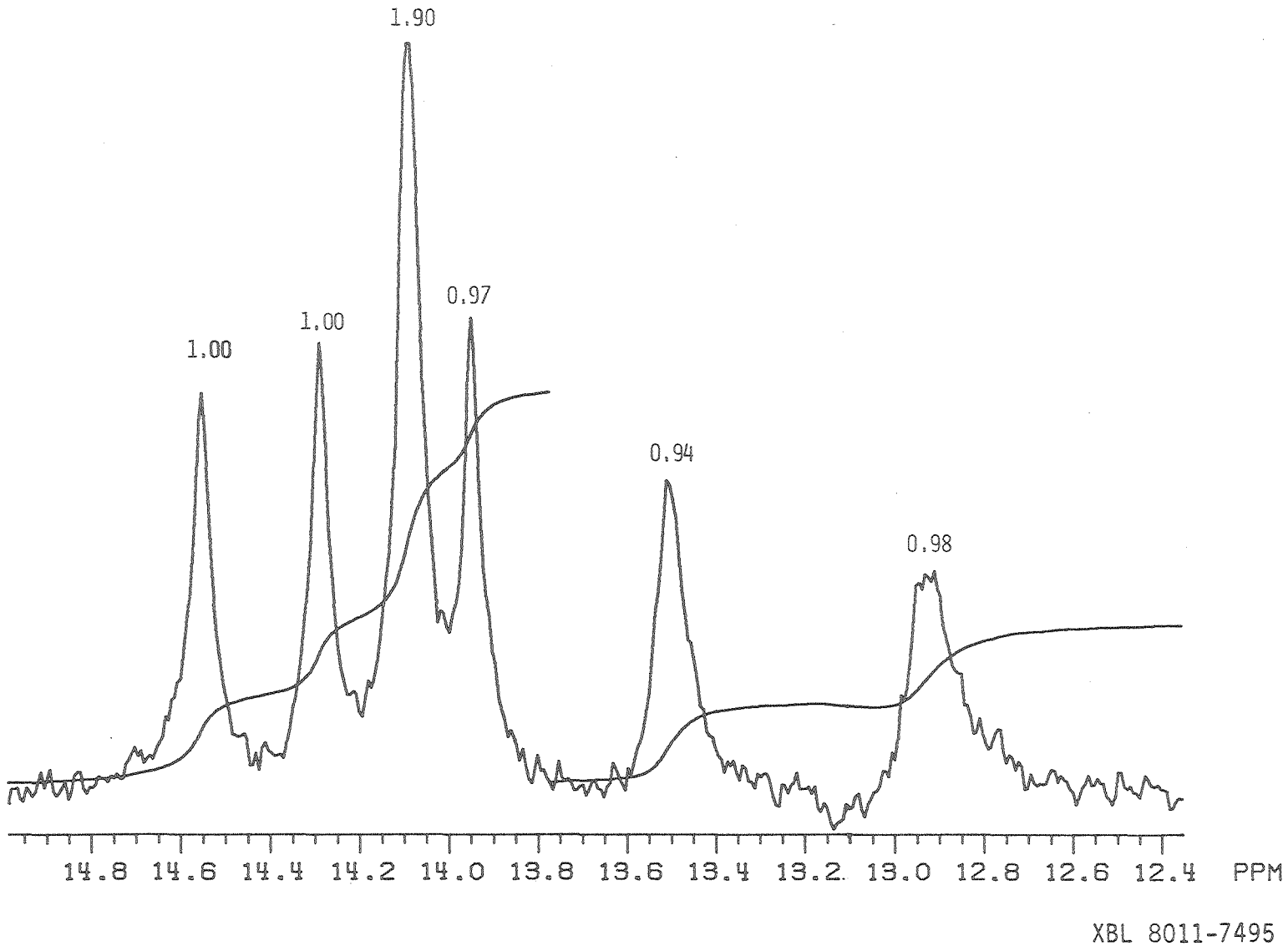
c) rCA_5G + dCT_5G double strand at 5°C.

d) dCA_5G + rCU_5G in triple strand at 1°C and in 1.0 M NaCl instead of 0.18 M NaCl.



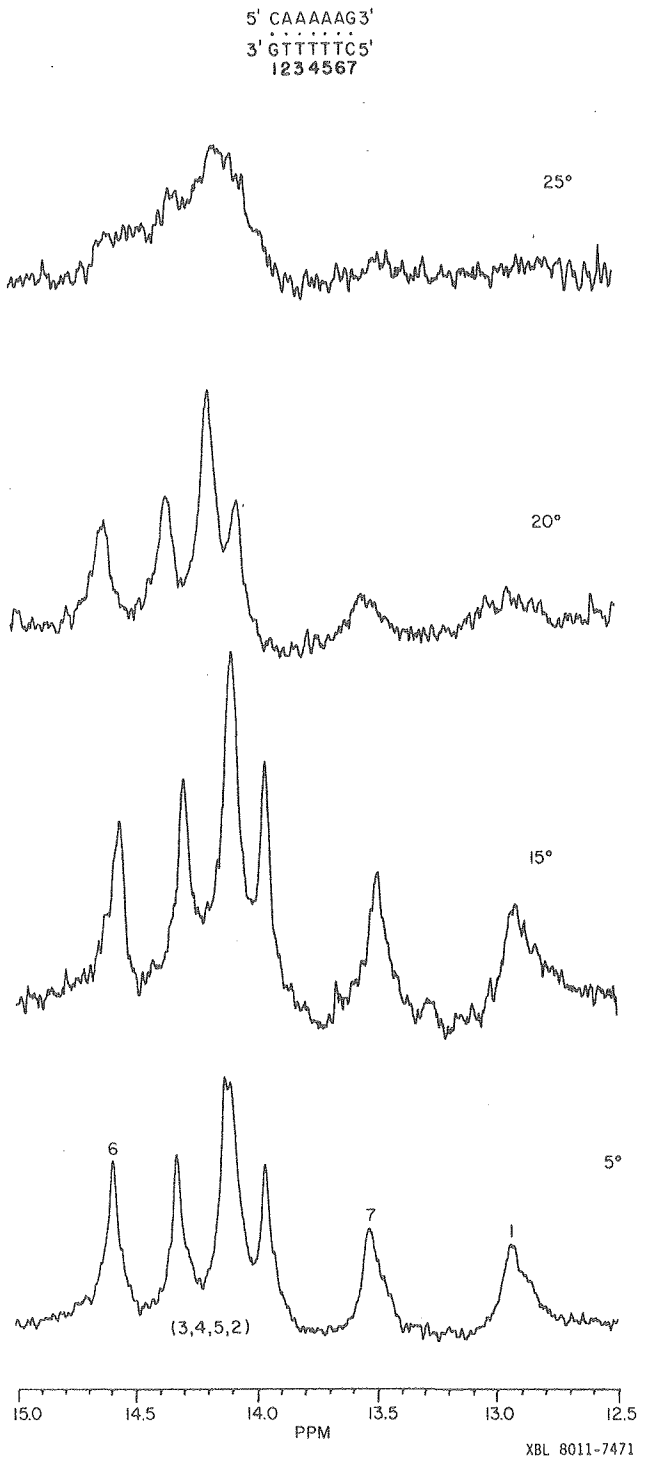
XBL 7911-13518

Figure 2.8. Integration of imino proton region of dCA_5G + dCT_5G helix.



XBL 8011-7495

Figure 2.9. Melting of imino proton region of $dCA_5G + dCT_5G$ helix.

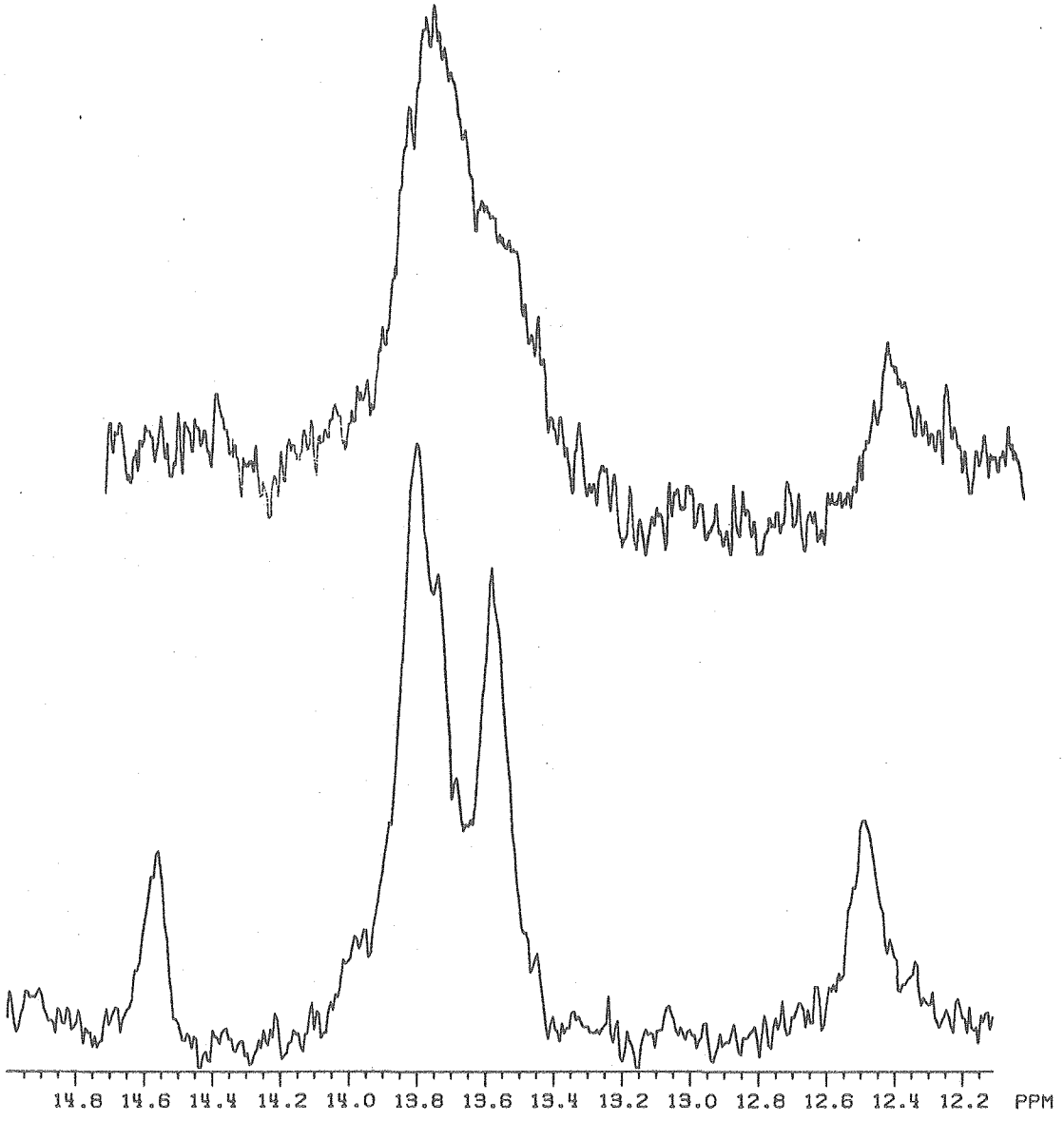

5' CAAAAAG3'
3' GTTTTTC5'
1234567
25°
20°
15°
5°
6
7
1
(3,4,5,2)
15.0
14.5
14.0
13.5
13.0
12.5
PPM
XBL 8011-7471

relative to the rest of the molecule, and by 25°C they have disappeared. The differential broadening observed is due to the faster exchange of the terminal base pair protons compared to those in the interior of the helix.[27,40] Calculated chemical shifts of the two C·G resonances allow assignment of the resonance at 12.9 ppm to the C·G in position one and the 13.55 ppm peak to the C·G in position seven (see Table VII).[41] The assignments of the A·T base pairs were made by comparison with the calculated shifts for DNA B geometry. This enabled us to assign the A·T resonances at the 6 position to 14.56 ppm with the rest of the A·T base pairs assigned to region 13.9 - 14.4 ppm.

In the rCU_5G + rCA_5G system we were aided in the assignments by comparison with the system rCA_6 + rCU_5G. Figure 2.10 shows the comparison of the rCA_5G + rCU_5G and rCA_6 + rCU_5G. This gives unambiguous assignments of the C·G(1) base pair to 12.6 ppm, A·U(6) to 14.6 ppm and C·G(7) to 13.6 ppm as shown in Figure 2.7. The other resonances in the rCU_5G + rCA_5G were assigned to the region 13.6 - 13.9 ppm. All chemical shifts at 5°C are given in Table VII.

In the duplex rCA_5G + dCT_5G, assignments were made in an analogous manner to the procedure used for the dCA_5G + dCT_5G helix. The C·G resonances again show very different temperature dependences than the A·T protons. The other peak assignments were then made using an RNA A geometry

Figure 2.10. Comparison of imino proton region of rCA_6 + rCU_5G (top) and rCA_5G + rCU_5G (bottom).



XBL 8011-7496

TABLE VII

EXPERIMENTAL AND CALCULATED CHEMICAL SHIFTS OF THE IMINO PROTONS IN THE DOUBLE HELIX

	#1 C G	#2 A T	#3 A T	#4 A T	#5 A T	#6 A T	#7 G C
dCA_5G + dCT_5G helix	12.94	(13.95	14.30	14.08	14.08)	14.56	13.51
rCA_5G + dCT_5G helix	12.38	13.40	(14.02	13.88	13.80)	14.53	13.52
rCA_5G + rCU_5G helix	12.52	(13.60	13.76	13.82	13.82)	14.60	13.60
DNA B form calculation	12.53	13.50	13.80	13.80	13.85	14.15	13.30
RNA A form calculation	12.80	13.46	13.70	13.70	13.70	14.10	13.50
RNA A' form calculation	12.72	13.27	13.58	13.58	13.60	14.11	13.47

The calculations were done using numbers derived by Arter and Schmidt,[41] and we assumed thymine has the same ring current as uracil. Values in parentheses indicate that we were unable to assign these peaks to particular bases.

to calculate ring current shifts for each resonance. For this hybrid the calculated ring current shifts are shown for RNA A and A' as well as DNA B form geometry in Table VII.

C) Temperature Dependence of the Nonexchangeable Protons in the Single Strands

The temperature dependences of the base protons of the single strands dCA_5G, dCT_5G, and rCA_5G are shown in Figures 2.11, 2.12, and 2.13, respectively. The cytosine H6 and H5, the adenine H2 and H8, the guanine H8 and the thymine H6 and methyl proton resonances were observed. The chemical shifts for all the base protons in dCT_5G show little change with temperature (< 0.1 ppm from 5 - 65°C, see Figure 2.12). The largest changes are on the adenine H2 protons in dCA_5G and the adenine H8 protons on rCA_5G as shown in Figures 2.11 and 2.13. Most of the proton chemical shift changes observed are not linear with temperature; this is important to note when analyzing melting curves, as will be discussed in a later section. There are large differences in chemical shift changes with temperature between dCA_5G and rCA_5G. The temperature dependences of the chemical shift for a particular proton are also sometimes in opposite directions for the two strands. For example, compare the H6 of C(1) and the H8 of A (3, 4, 5 or 6) on the two strands in Figure 2.14. These observations indicate different conformations for the deoxyribo- and ribo- strands of CA_5G.

Figure 2.11. Temperature dependence of the base protons in the single strand of dCA_5G.

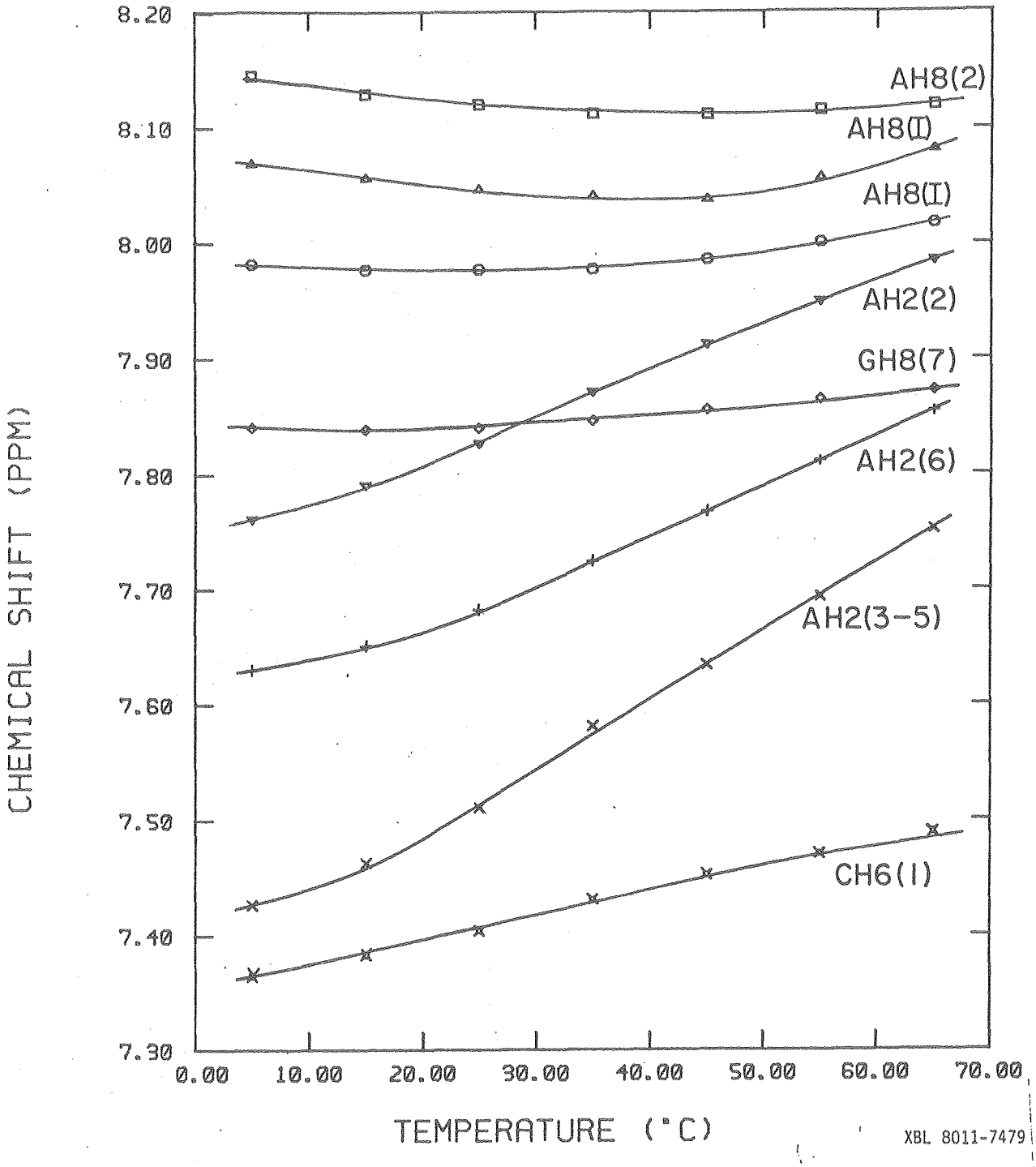

8.20
8.10
8.00
7.90
7.80
7.70
7.60
7.50
7.40
7.30
CHEMICAL SHIFT (PPM)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
AH8(2)
AH8(I)
AH8(I)
AH2(2)
GH8(7)
AH2(6)
AH2(3-5)
CH6(1)
XBL 8011-7479

Figure 2.12. Temperature dependence of base protons in the single strand of dCT_5G.

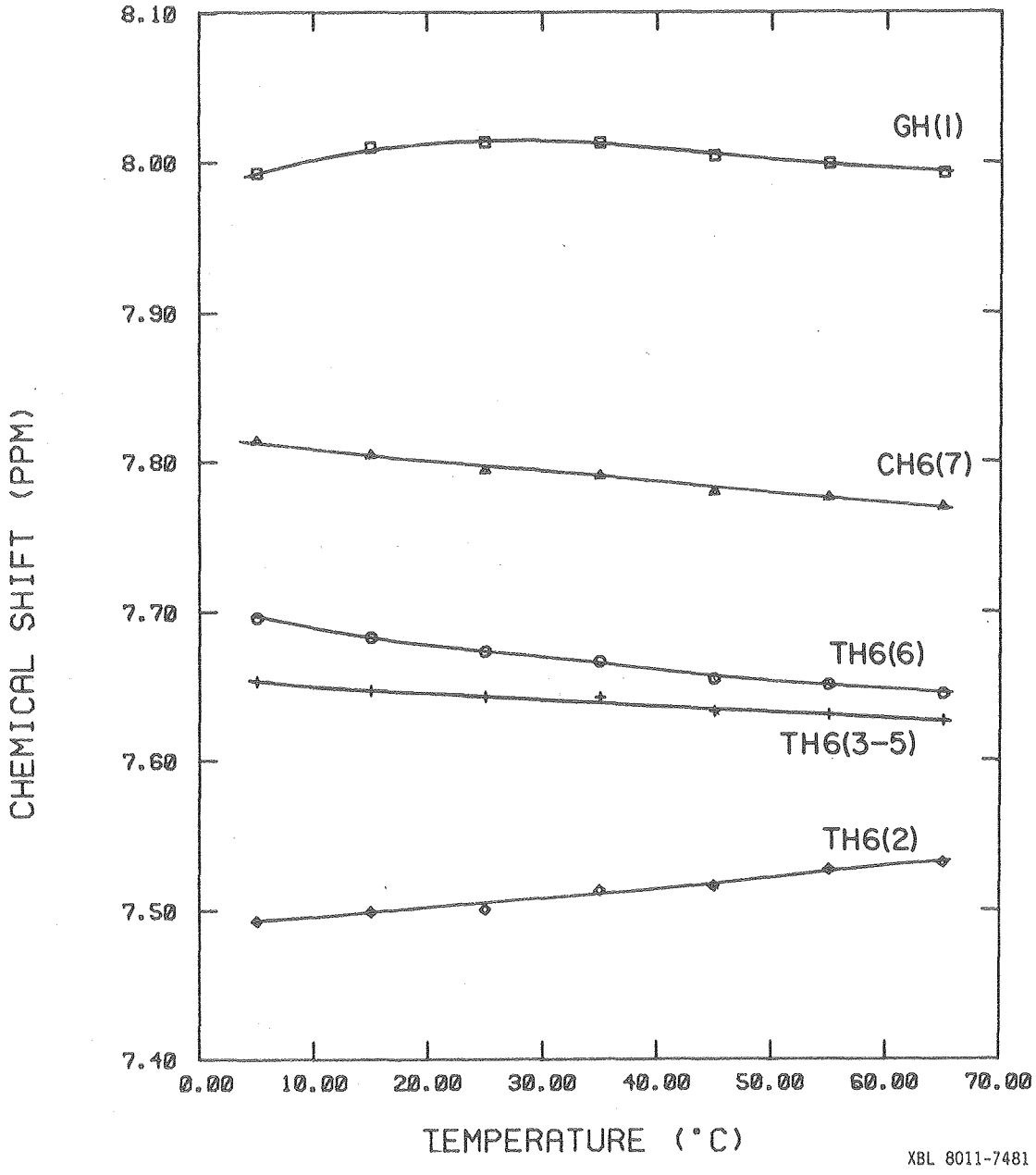
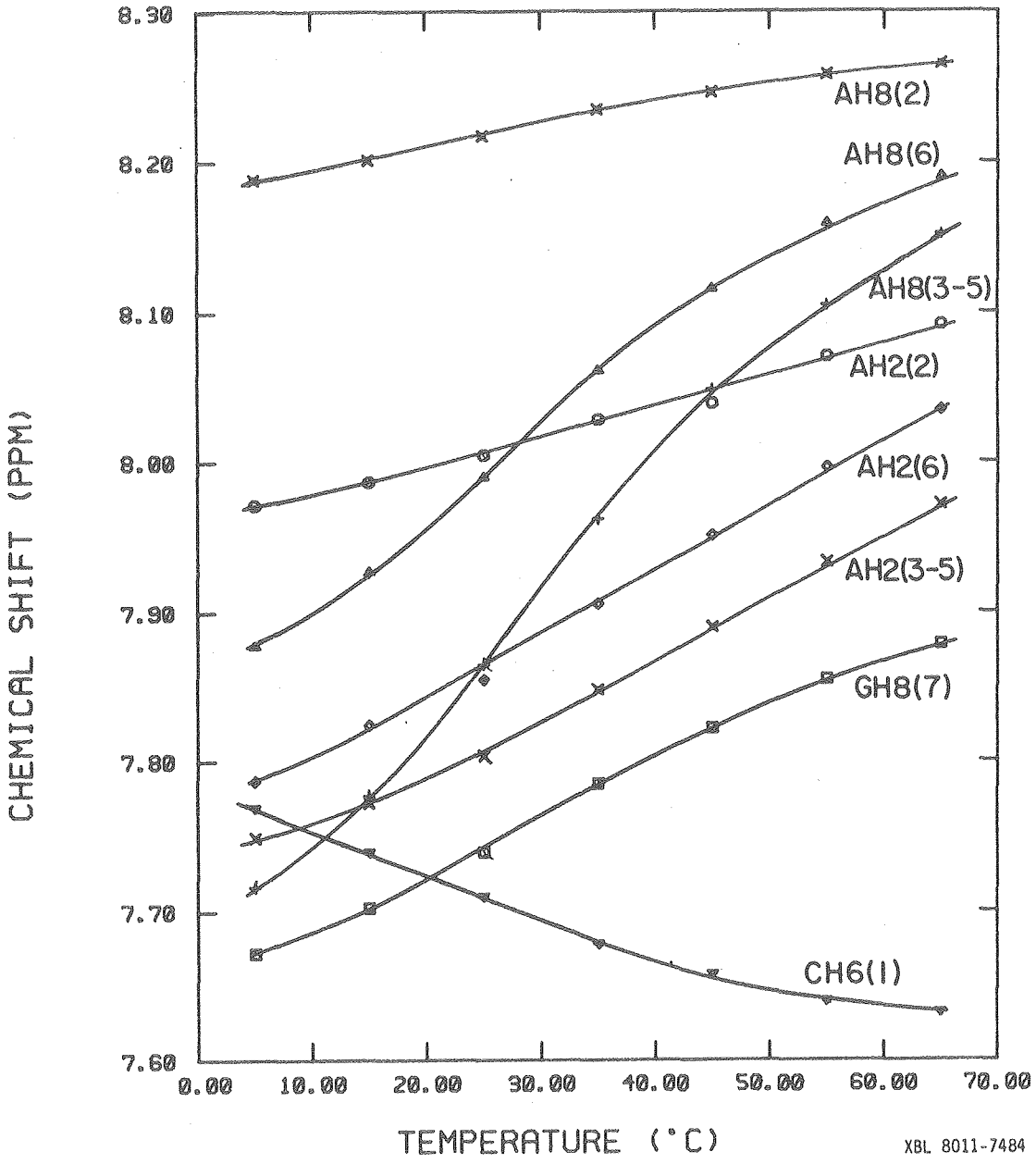

8.10
8.00
7.90
7.80
7.70
7.60
7.50
7.40
CHEMICAL SHIFT (PPM)
GH(1)
CH6(7)
TH6(6)
TH6(3-5)
TH6(2)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
XBL 8011-7481

Figure 2.13. Temperature dependence of base protons in the single strand of rCA_5G.

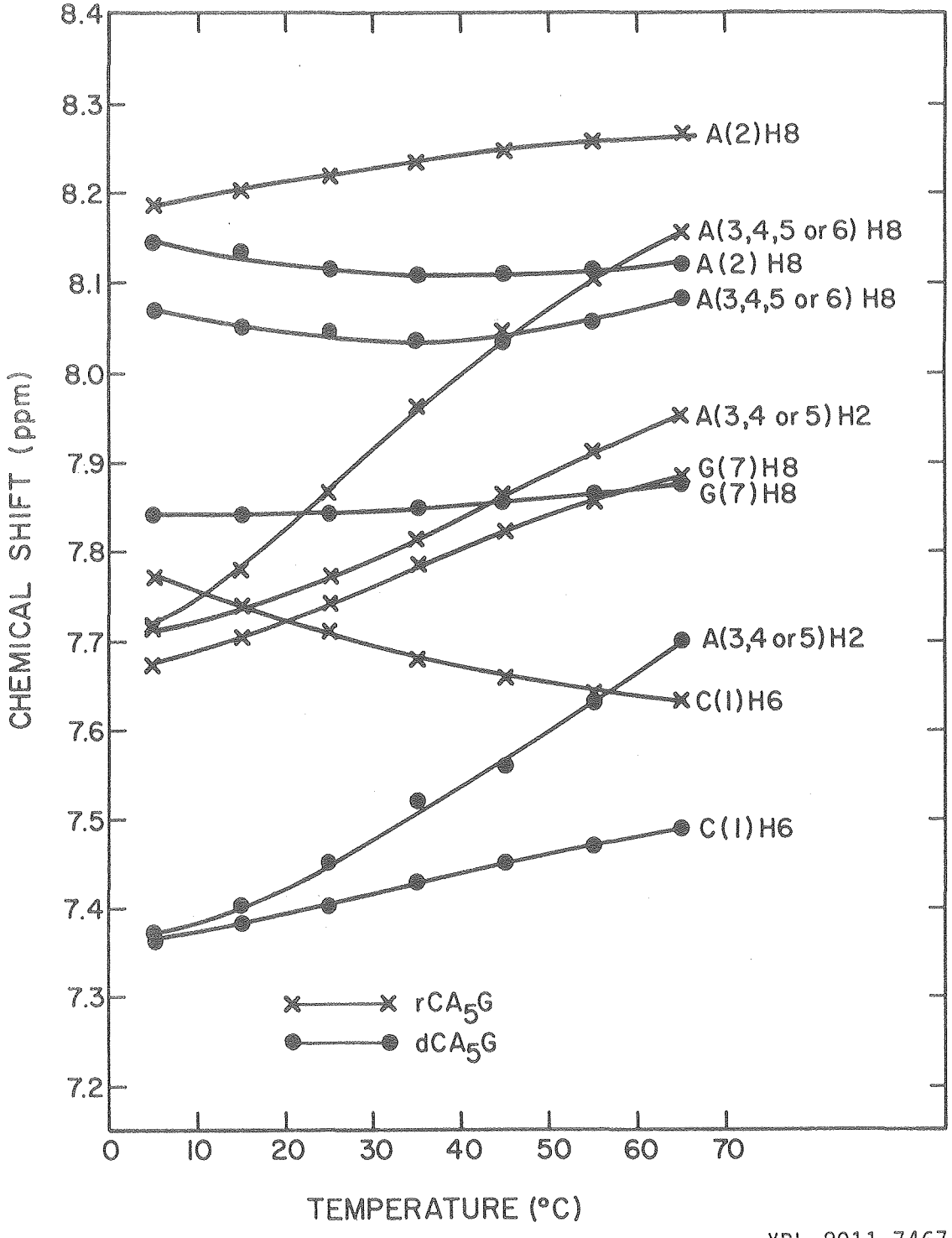

8.30
8.20
8.10
8.00
7.90
7.80
7.70
7.60
CHEMICAL SHIFT (PPM)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
AH8(2)
AH8(6)
AH8(3-5)
AH2(2)
AH2(6)
AH2(3-5)
GH8(7)
CH6(1)
XBL 8011-7484

The best measures of base-base stacking from chemical shifts are obtained from the H2 of adenine and the H5 of the pyrimidines.[42] The H8 of adenines and the H6 of pyrimidines have significant effects basides ring currents, which effect their chemical shifts,[42,43] such as the glycosidic torsion angle, and the proximity of the sugar ring oxygens or phosphate groups. Therefore, in order to get an idea of the relative base-base stacking in the dCA_5G and rCA_5G strands, we compared the chemical shifts of the adenine H2 protons. The adenine H2 protons on the ribose strand are always downfield (less shielded) from the same protons on the deoxyribose strand. This is indicative of more base-base stacking in dCA_5G than in rCA_5G.

D) The Helix to Coil Transition of dCT_5G + dCA_5G and dCT_5G + rCA_5G as Followed by the Chemical Shift of the Base Protons

The temperature dependence of the chemical shifts of the base protons were studied in the oligonucleotides, dCT_5G + dCA_5G, and dCT_5G + rCA_5G. In the helix-to-coil transition of the dCA_5G + dCT_5G, all the base proton chemical shifts were followed from 5 - 65°C. In the intermediate states from 20 - 35°C, it is difficult to obtain exact chemical shifts of some of the protons due to extensive broadening of the resonances undergoing chemical exchange.

Figure 2.14. Comparison of temperature dependence of base protons in dCA_5G and rCA_5G.



XBL 8011-7467

In the other helix, $dCT_5C + rCA_5G$, the chemical exchange broadening over the 20 - 30°C temperature range prevented the assignment of many of the resonances at lower temperatures. The chemical shift changes with temperature for the base protons in the two helices are shown in Figures 2.15a-e and 2.16a and b.

In order to monitor the helix-to-coil transitions in these oligomers, base protons were chosen which show large changes upon formation of the helical state (> 0.1 ppm). The curves for many of these protons have sigmoidal shapes and characteristics which are very similar to absorbance versus temperature curves for these oligomers.[44] However, the NMR data allows one to follow the properties of individual bases throughout the double-to-single strand transition. Thus NMR is more useful than absorbance studies for finding properties such as differential melting of the helix.[15,27,40]

The T_m's or melting temperatures of different bases were compared to learn if there is any melting on the ends of the $dCT_5G + dCA_5G$ helix. The analysis was the same as that used in optical studies;[25] this requires the melting behavior of the single strand to obtain the T_m. We measure the chemical shifts versus temperature of the single strands directly by using solutions containing only one of the strands. The double strand baseline can only be estimated from the low temperature behaivor of the chemical shifts. We used the

Figure 2.15. Temperature dependence of the following protons in the helix-to-coil transition of $dCA_5G + dCT_5G$.

a) Thymine H6 protons.

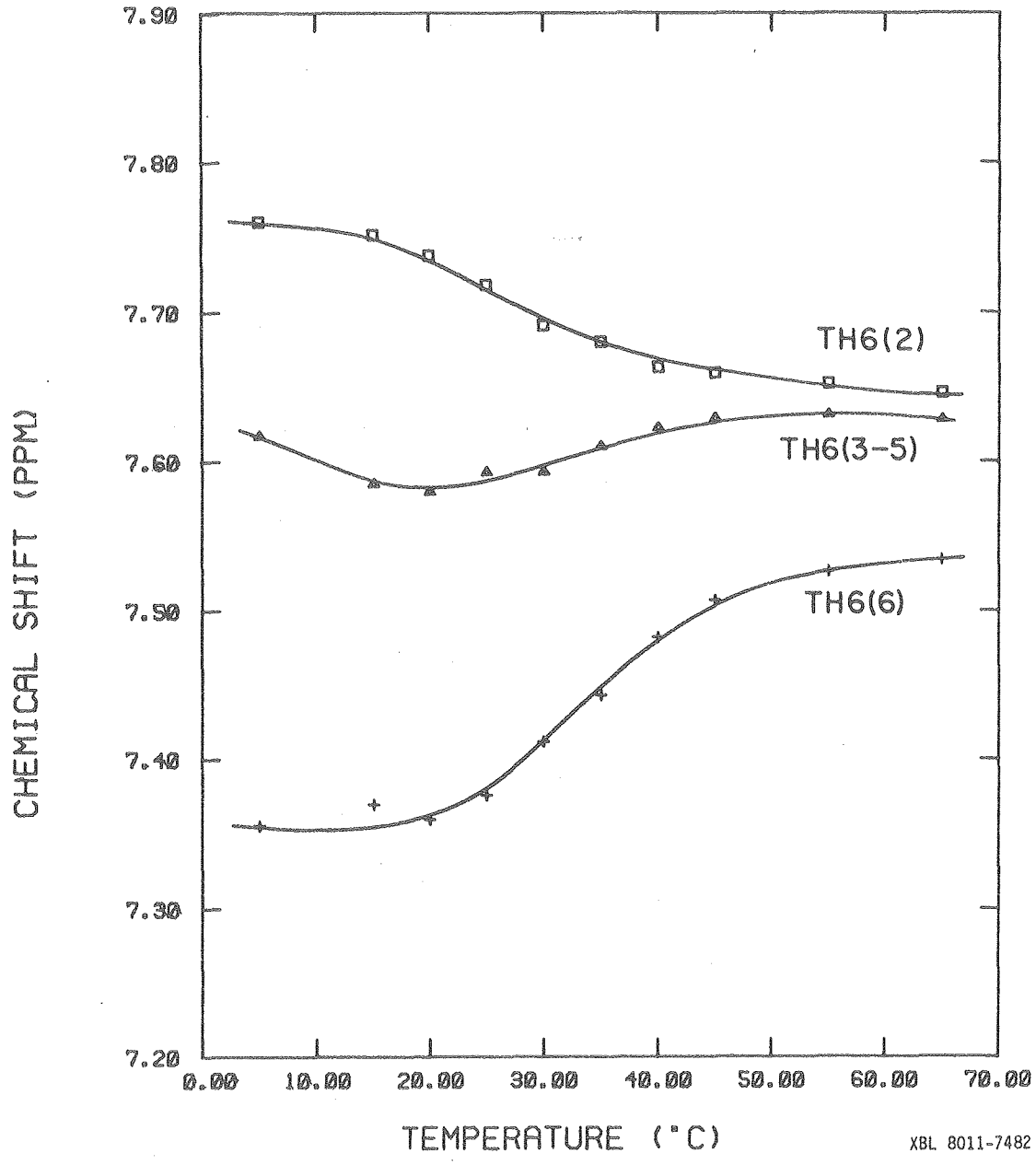
b) Terminal guanine H8 and cytosine H6 protons.

c) Adenine H8 protons.

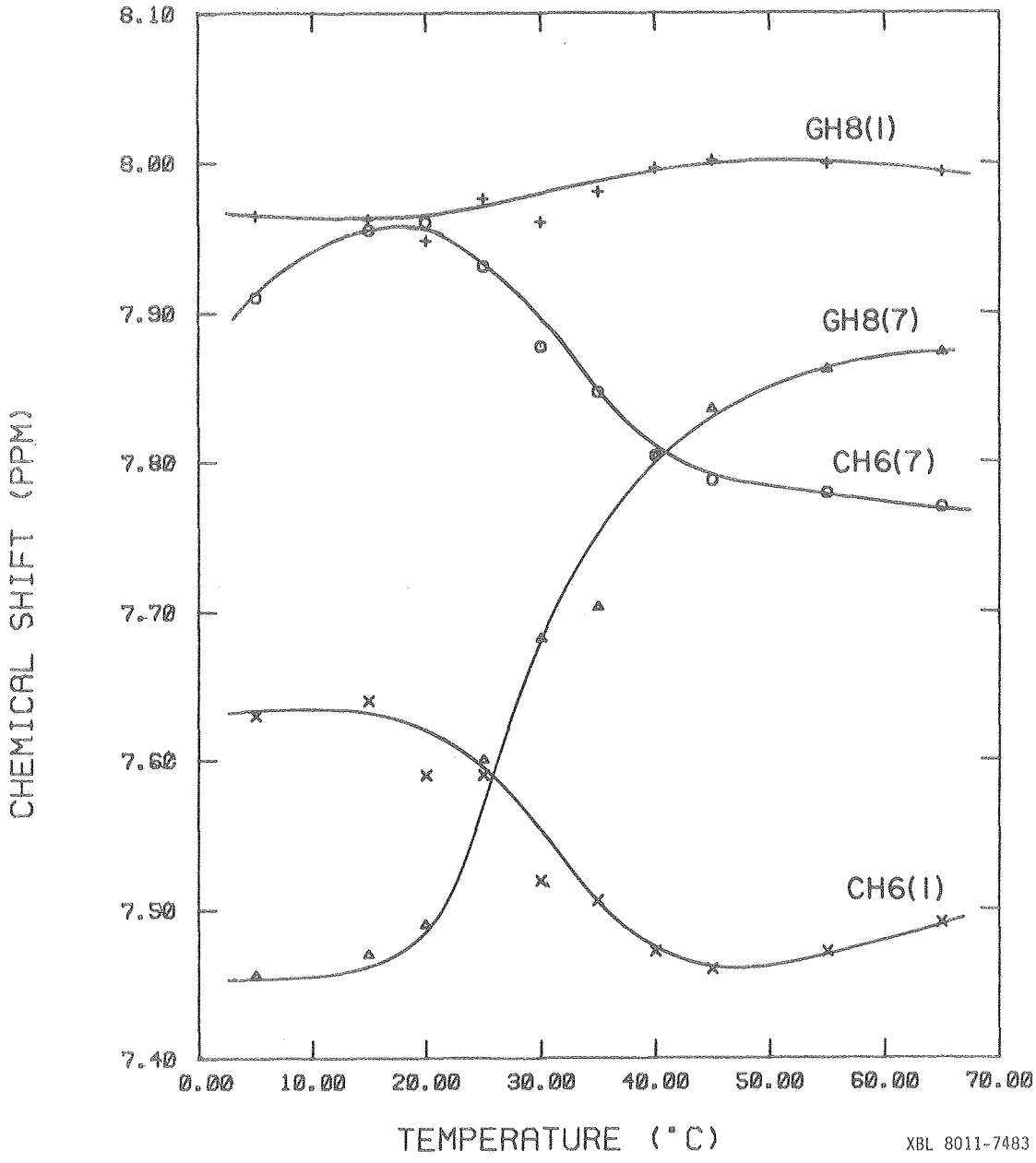
d) Adenine H2 protons.

e) Thymine methyl protons.

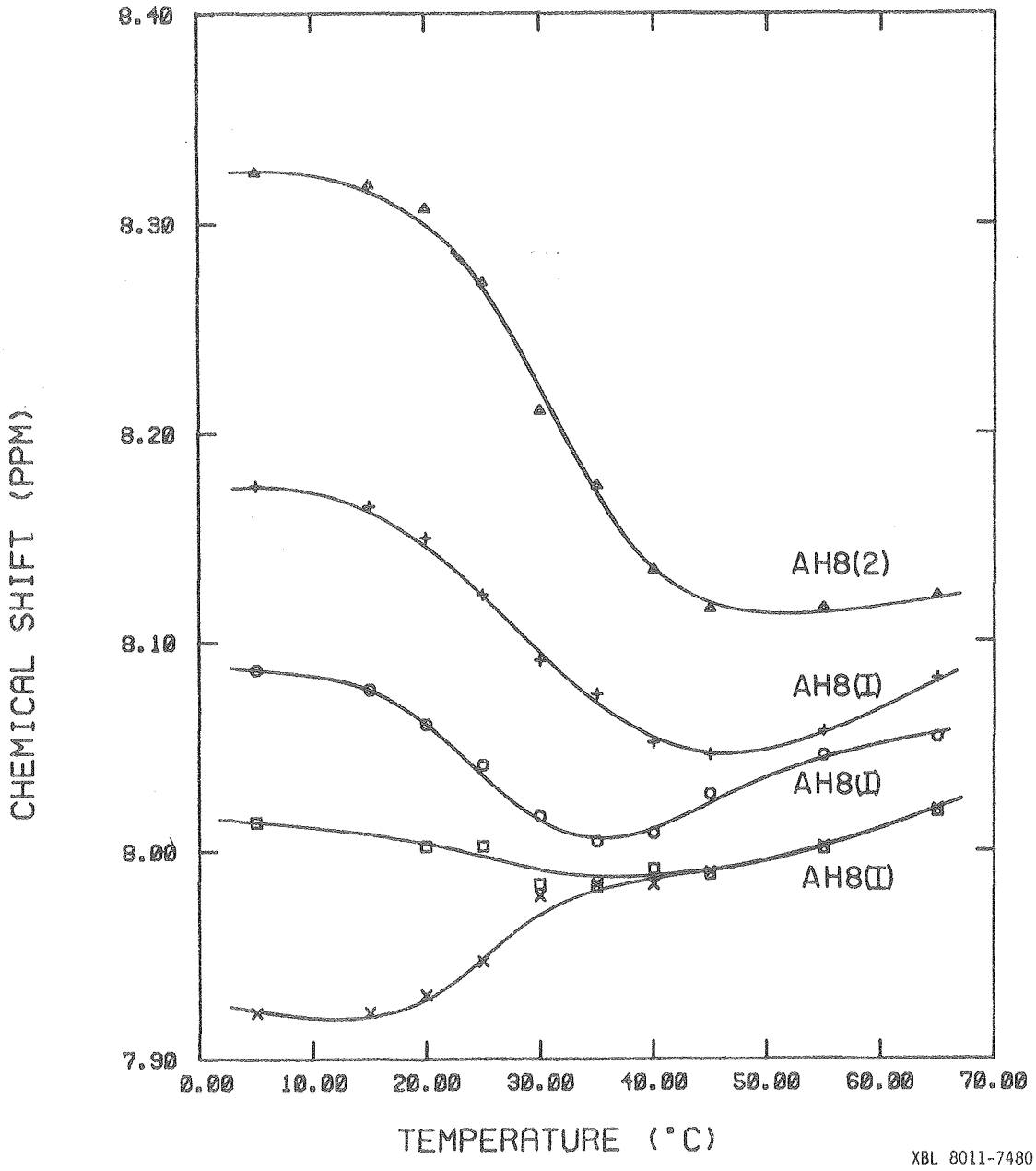
a)


7.90
7.80
7.70
7.60
7.50
7.40
7.30
7.20
CHEMICAL SHIFT (PPM)
TH6(2)
TH6(3-5)
TH6(6)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
XBL 8011-7482

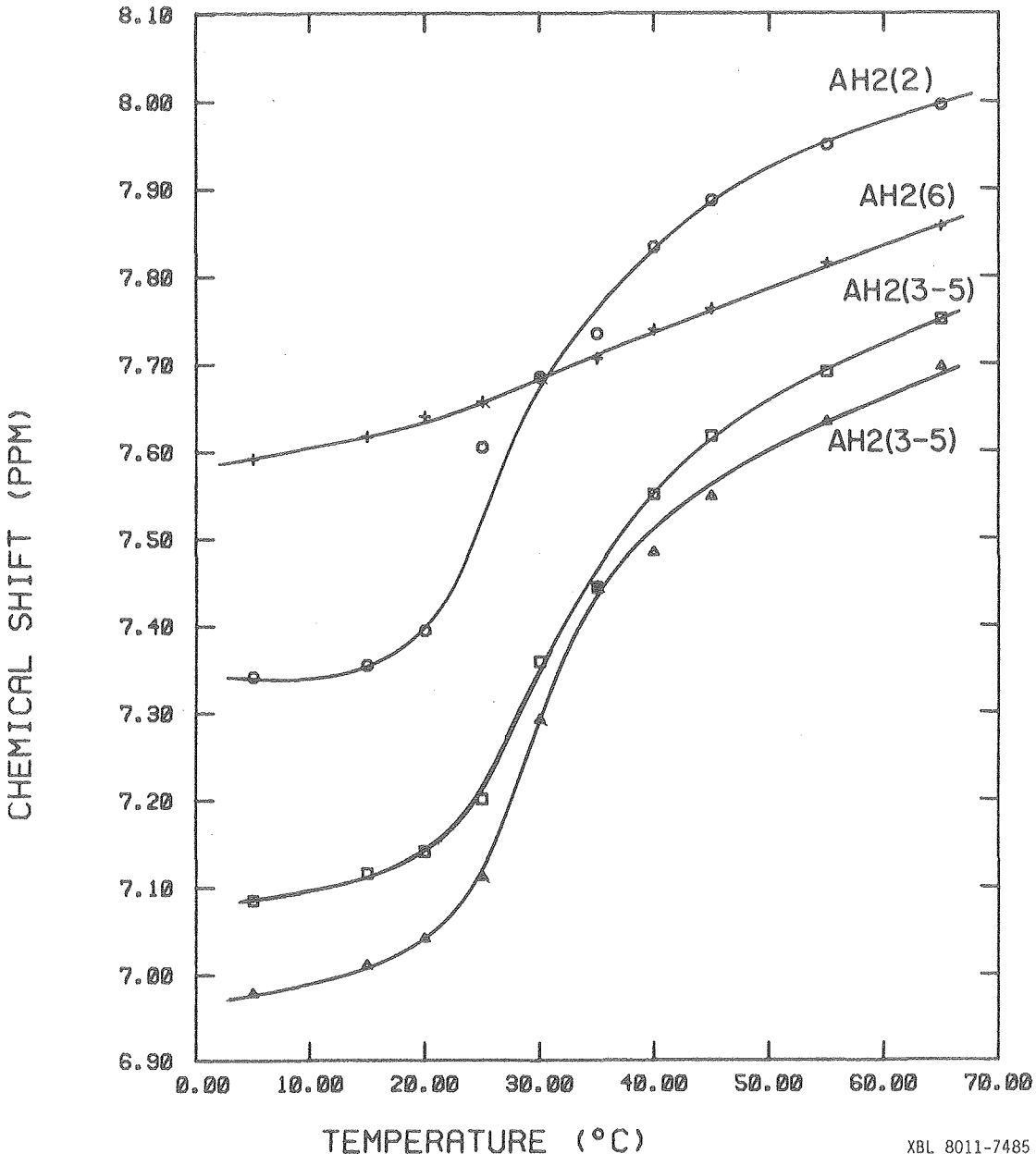
b)


8.10
8.00
7.90
7.80
7.70
7.60
7.50
7.40
CHEMICAL SHIFT (PPM)
GH8(I)
GH8(7)
CH6(7)
CH6(I)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
XBL 8011-7483

c)


8.40
8.30
8.20
8.10
8.00
7.90
CHEMICAL SHIFT (PPM)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
AH8(2)
AH8(I)
AH8(I)
AH8(I)
XBL 8011-7480

d)


8.10
8.00
7.90
7.80
7.70
7.60
7.50
7.40
7.30
7.20
7.10
7.00
6.90
CHEMICAL SHIFT (PPM)
AH2(2)
AH2(6)
AH2(3-5)
AH2(3-5)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
XBL 8011-7485

e)

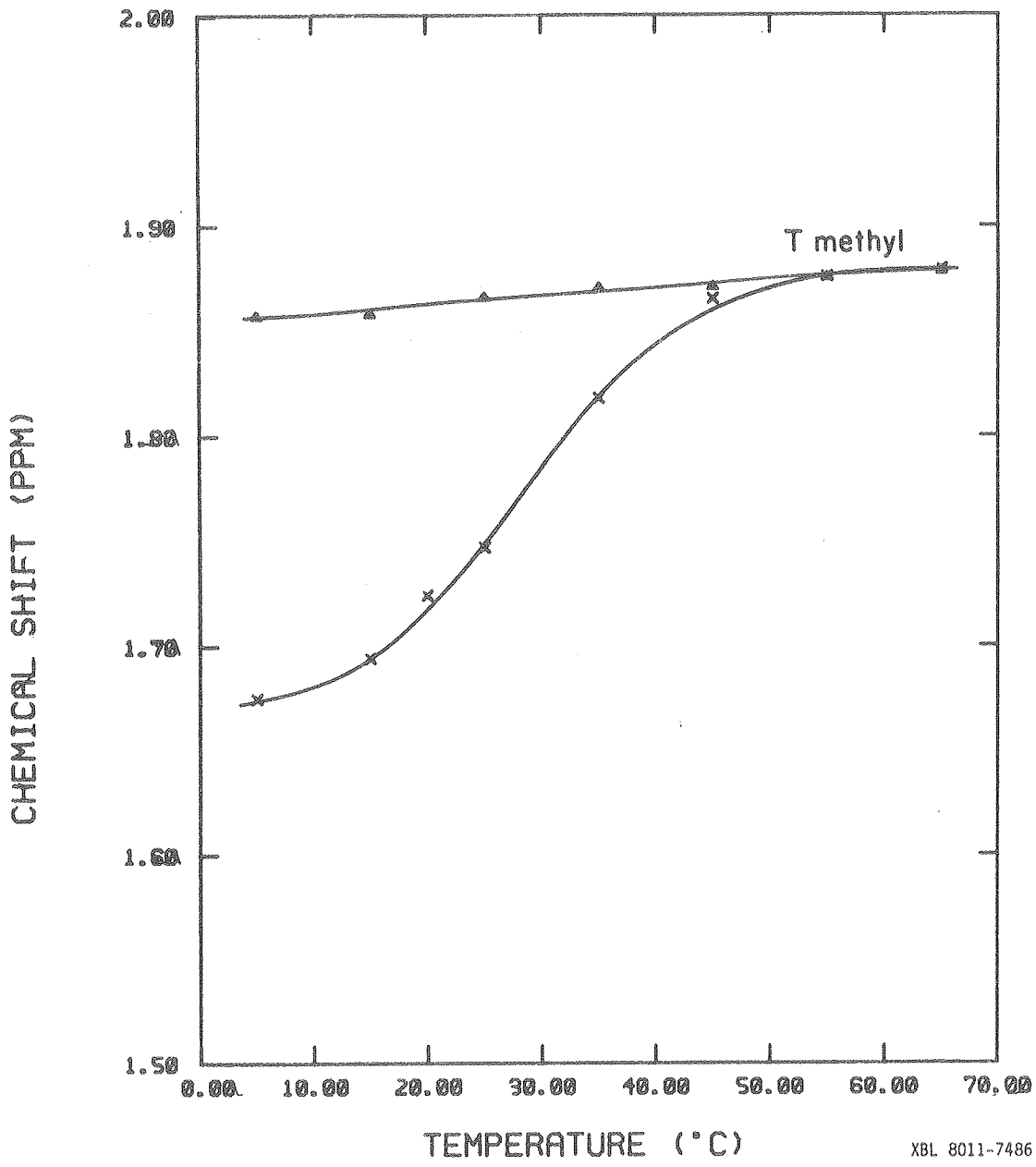
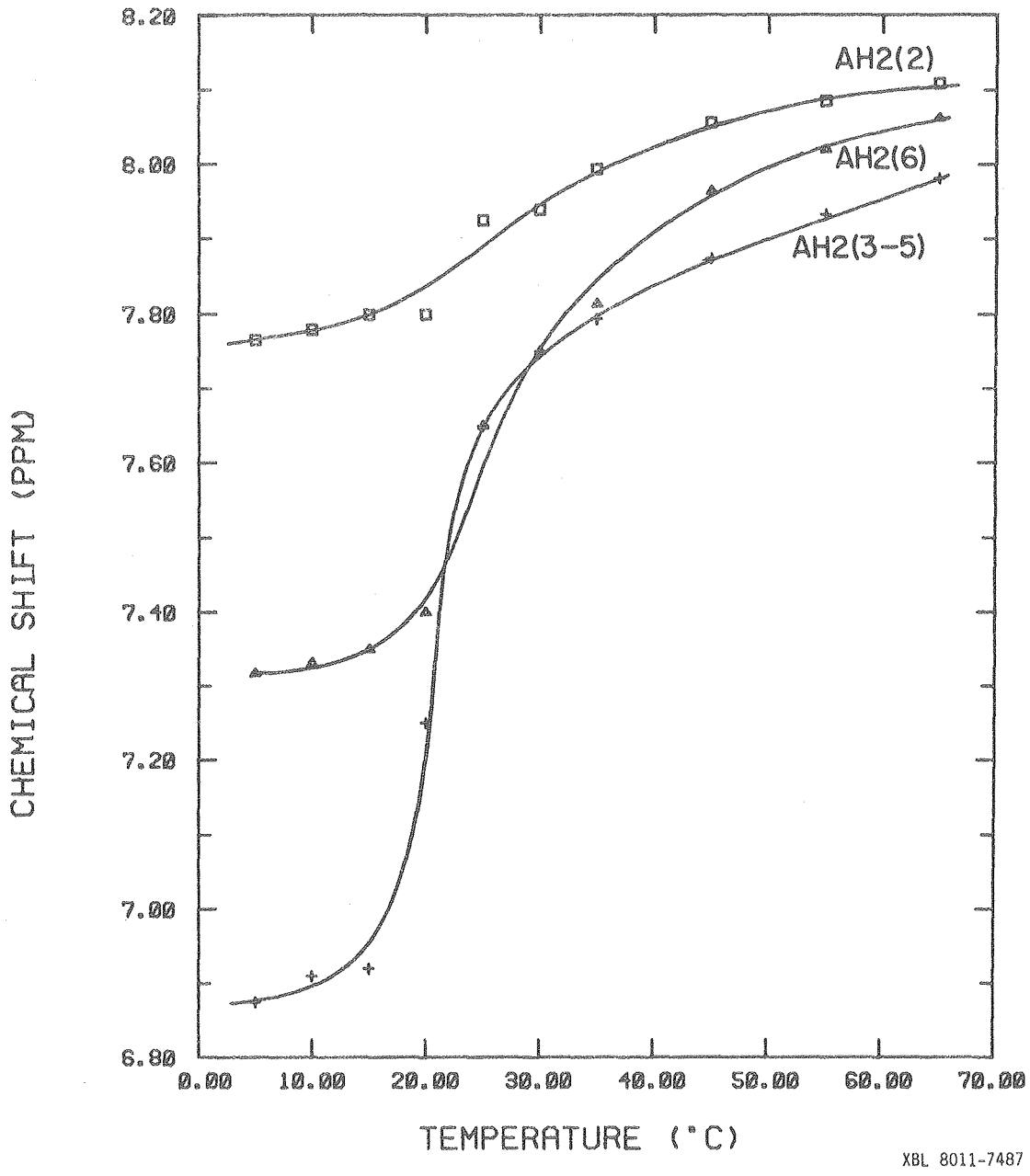

2.00
1.90
1.80
1.70
1.60
1.50
CHEMICAL SHIFT (PPM)
T methyl
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE (°C)
XBL 8011-7486

Figure 2.16. Temperature dependence of the following protons during the helix-to-coil transition of $dCT_5G + rCA_5G$.

a) Adenine H2 protons.

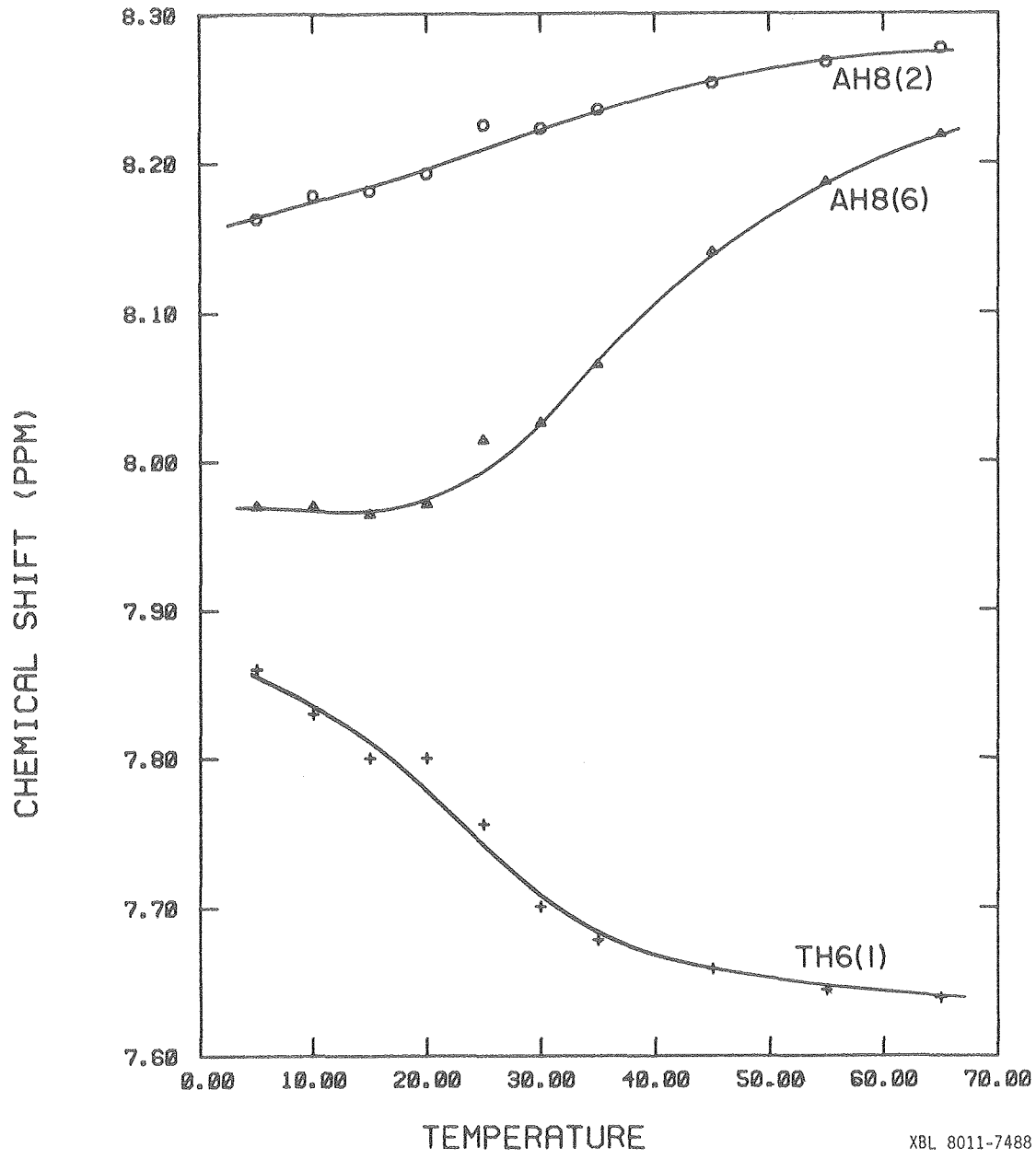
b) Adenine H8 and thymine H6 protons.

a)



XBL 8011-7487

b)


8.30
8.20
8.10
8.00
7.90
7.80
7.70
7.60
CHEMICAL SHIFT (PPM)
AH8(2)
AH8(6)
TH6(1)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
TEMPERATURE
XBL 8011-7488

value of the chemical shift measured at 5°C for the lower baseline. For most of the protons there are very small chemical shift changes between 5 and 15°C so that the chemical shift of the completely double strand was taken to be temperature independent. The T_m values are all very similar for protons in different parts of the helix, with an average value of 29 ± 2°C at a concentration of 1.0 mM per strand. Since all the base protons melt at approximately the same temperature, the concentration of partially formed helices is small.

Optical melting under identical conditions gives a T_m of 34 ± 2°C. No change in the melt was seen whether H_2O or D_2O was used as solvent. The difference in the T_m's obtained by NMR and optically is probably due to assumptions about the temperature dependence of the properties of the double strand, and also errors involved in assuming fast exchange, on the NMR time scale, throughout the single to double strand transition. These effects will be discussed in a later section.

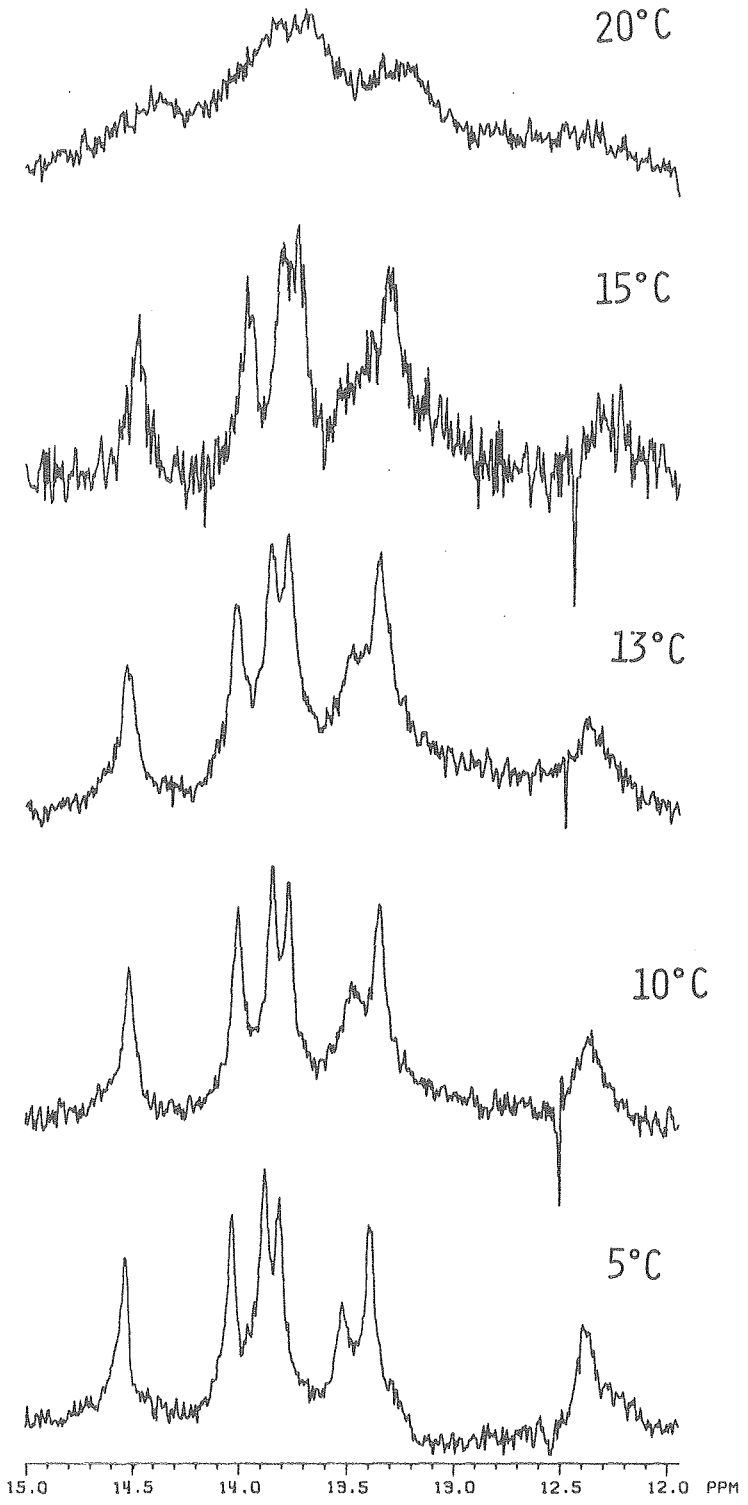
Melting curves for the oligomer $dCT_5G + rCA_5G$ are shown in Figures 2.16a and b. The temperature dependence of only a small number of the base protons are shown since exchange broadening and overlap of too many protons hindered the identification of many of the protons in the double strand. In this helix we were unable to follow any of the guanine

or cytosine protons throughout the whole melting transition, so that it was not possible to tell if there was significant melting of the ends of the helix. The average T_m for all the base pairs was 23.5°C at 0.5 mM per strand. The T_m obtained under similar conditions from optical data extrapolated to the same concentration was 25.7°C.[44]

E) Temperature Dependence of the Exchangeable Imino Protons

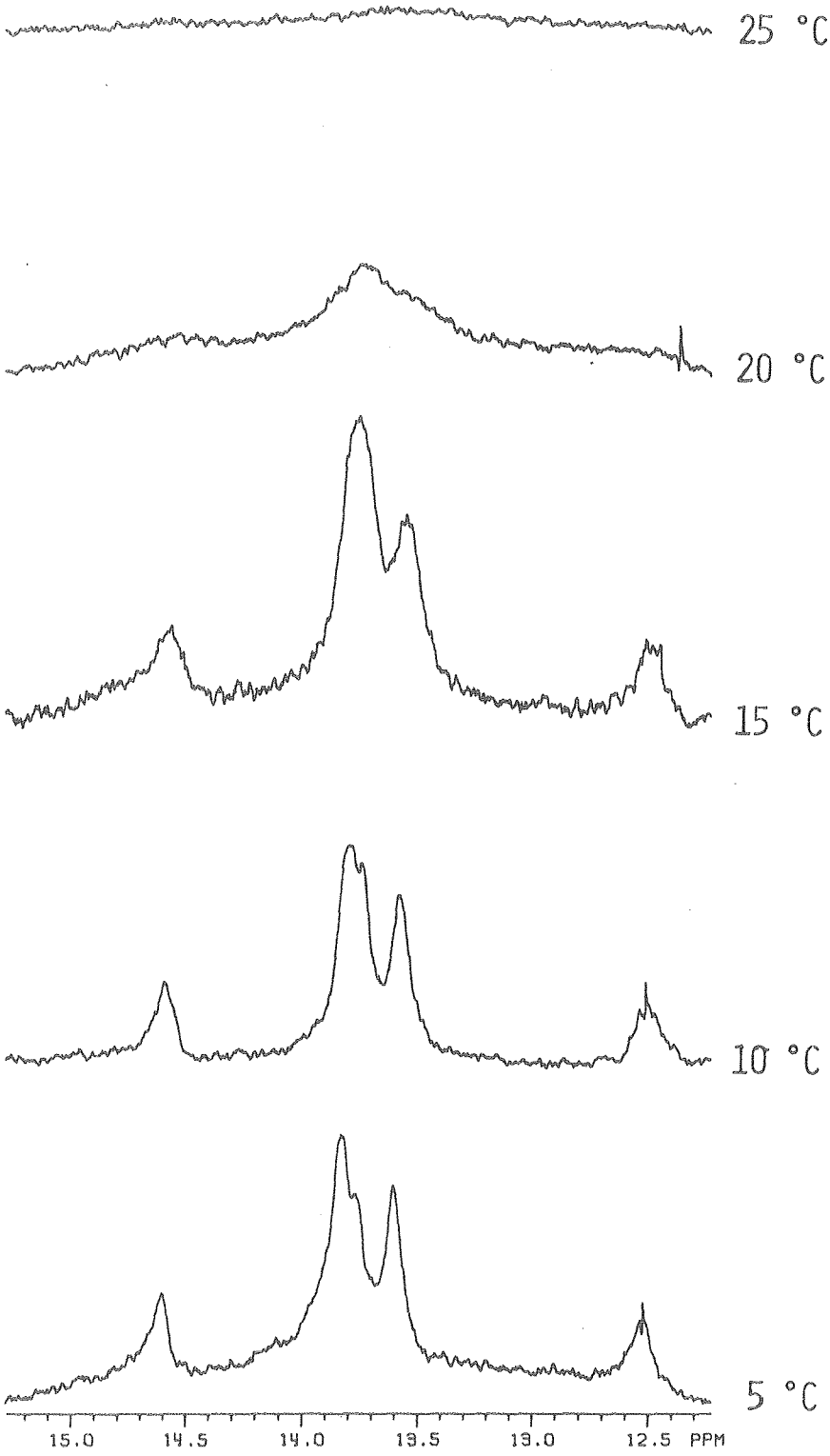
The temperature dependences of the imino base protons in dCT_5G + dCA_5G are shown in Figure 2.9. As already discussed in the section on the assignments of the imino protons, it is found that the base pairs on the ends broaden and disappear before the rest of the base pairs in the helix. Similar broadening is observed for the other helices dCT_5G + rCA_5G, and rCU_5G + rCA_5G and are shown in Figures 2.17 and 2.18. The broadening phenomenon observed is due to chemical exchange as will be discussed in a later section.[45,46] The C·G imino protons for the three systems broaden and disappear in the range from 17 - 23°C. There is little chemical shift change with temperature for any of these resonances. Differences in the extent of broadening for a particular proton at a given temperature are seen in the three systems; this is due to different lifetimes for exchange of the imino protons with H_2O. This will be discussed in a later section.

Figure 2.17. Melting of imino protons in rCA_5G + dCT_5G helix.



XBL 8011-7493

Figure 2.18. Melting of imino protons in rCA_5G + rCU_5G helix.



XBL 8011-7489

F) Sugar Pucker of the Furanose Ring in the Double Strand

In two of the oligomer systems, the sugar pucker for some of the ribose or deoxyribose rings were calculated. The conformation of the ribose or deoxyribose ring has been described by Altona and Sundaralingam (1973) as a two-state equilibrium between type N[C2' exo, C3' endo] and type S[C2' endo, C3' exo]. Figure 2.19 shows an illustration of the different sugar conformations for 2' endo and 3' endo sugar puckers. The ribose ring conformation can then be calculated from the fact that in N type conformation J1'2' = ∿ 0 Hz and the S type has J1'2' = ∿ 10 Hz. For the deoxyribose ring the N type conformation has the property of J1'2' + J1'2" = ∿ 7 Hz while the S type has a value of ∿ 16 Hz for this sum.[47]

The assignment of the H1' protons to particular sugars in the oligomers was not possible, and coupling constants were obtained only for those resonances which were well separated from the main group of H1' protons.

Figure 2.20 shows the H1' deoxyribose proton region of dCT_5G + dCA_5G in the double strand at 5°C. Although the spectrum is fairly noisy, the two H1' peaks at 5.67 and 5.48 ppm had high enough signal to noise, and were well enough resolved, to do simulation. The simulation was done using a Nicolet 1180 computer in the NTCSIM routine. The lower part of the figure shows a blow-up of the simulated and experimental spectra. The best fit is obtained when

Figure 2.19. Illustration of the conformation of the sugar ring in 2' endo and 3' endo sugar pucker.

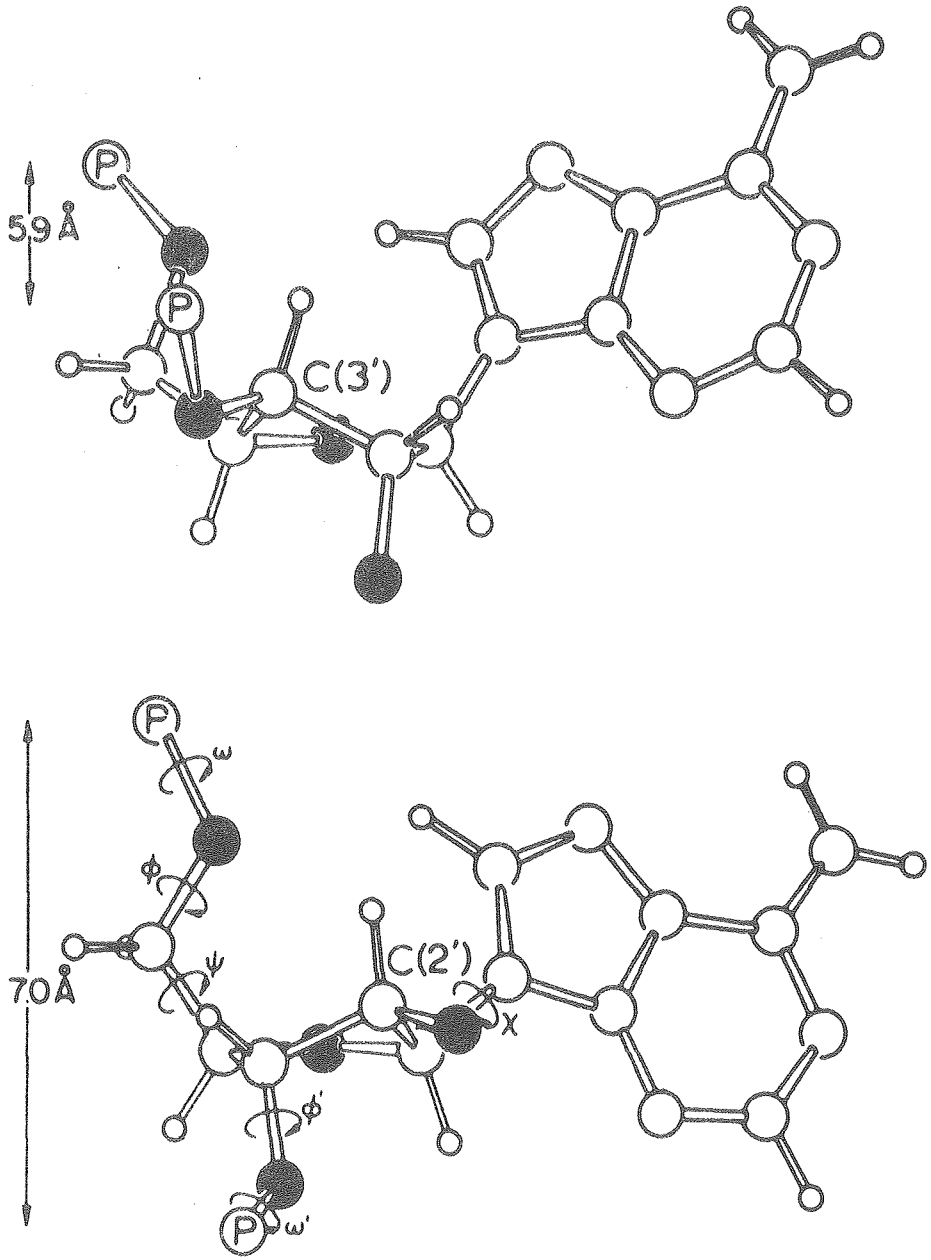
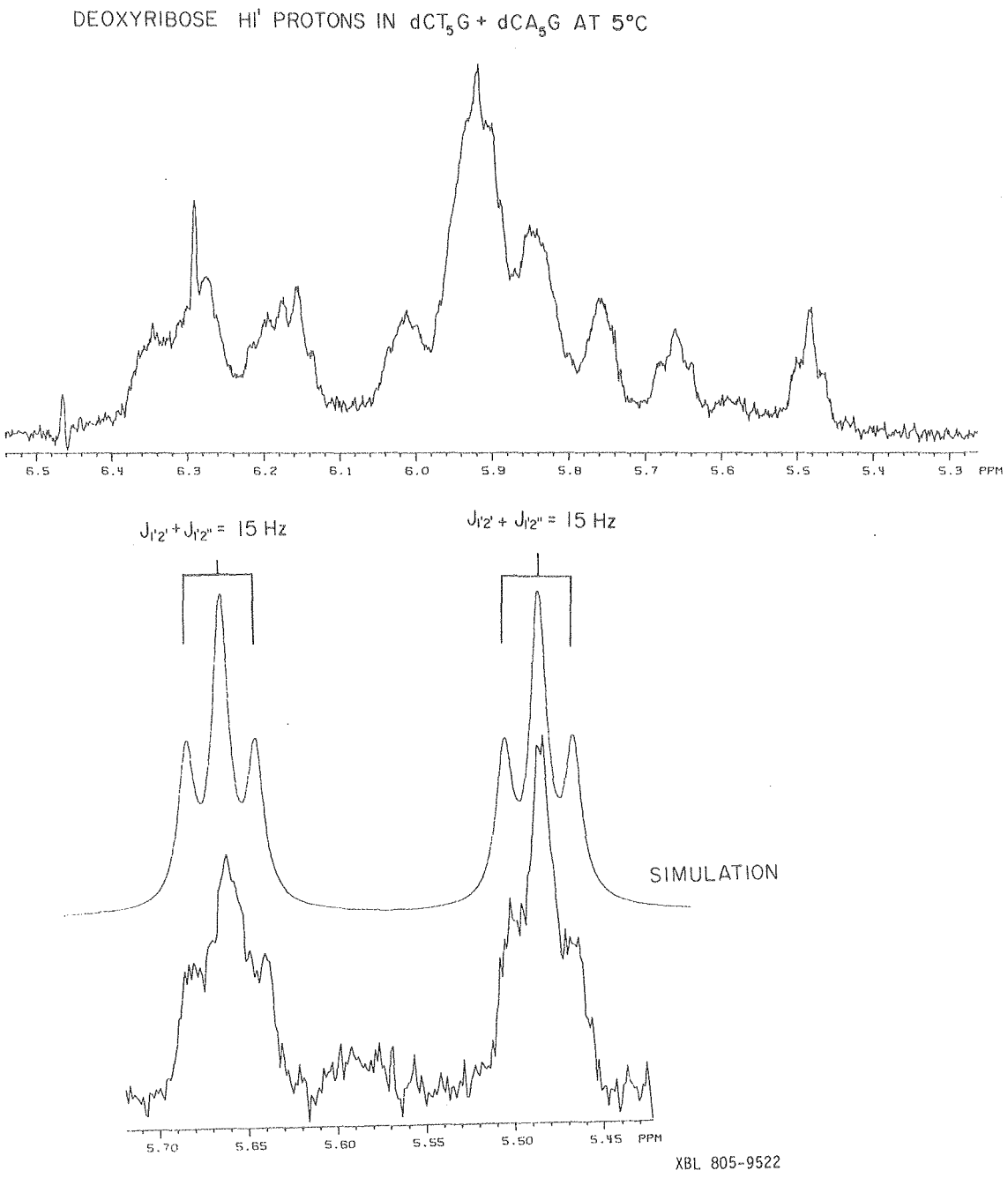


FIGURE 2. The two preferred or rigid nucleotide units. $C^{3'}$-*endo* (top), $C^{2'}$-*endo* (bottom). The torsion angle Ψ' is *gauche* in C(3′)-*endo* and *trans* in C(2′)-*endo*. (From Sundaralingam, M., *Structure and Conformation of Nucleic Acid and Protein-Nucleic Acid Interactions*, (Part 5), University Park Press, Baltimore, 1975, 487. With permission.)

XBL 8011-7466

Figure 2.20. Deoxyribose H1' protons in dCA_5G + dCT_5G double helix at 5°C.


DEOXYRIBOSE H1' PROTONS IN dCT5G + dCA5G AT 5°C
6.5 6.4 6.3 6.2 6.1 6.0 5.9 5.8 5.7 5.6 5.5 5.4 5.3 PPM
J1'2' + J1'2" = 15 Hz
J1'2' + J1'2" = 15 Hz
SIMULATION
5.70 5.65 5.60 5.55 5.50 5.45 PPM
XBL 805-9522

J1'2' + J1'2" = 15 ± 1 Hz. The sugar conformation is then calculated to be 90 ± 10% S type (2' endo). Although we were not able to uniquely assign these two H1' protons, the temperature dependence of these peaks, as well as comparison with the single strands at 65°C shows that both protons are from the dCA_5G strand. The chemical shifts of the H1' protons in the components of dCA_5G indicate that these two protons are from adenine residues. The fact that these deoxyribose sugars are in approximately 90% S type (2' endo) conformation is consistent with a B form type geometry for this helix.

Figure 2.21 shows the H1' ribose region of the double strand dCT_5G + rCA_5G at 5°C. This region contains only the H1' protons from the ribose CA_5G strand, as well as a peak at 5.92 ppm probably from a H1' deoxyribose proton. Simulation of this region shows that the J1'2' coupling constant is less than 1.5 Hz for all the H1' protons. This means that in the helical state there is > 85% N type (3' endo) sugar pucker in all the ribose rings in the rCA_5G strand. This 3' endo sugar pucker is indicative of an A type geometry for this helix.

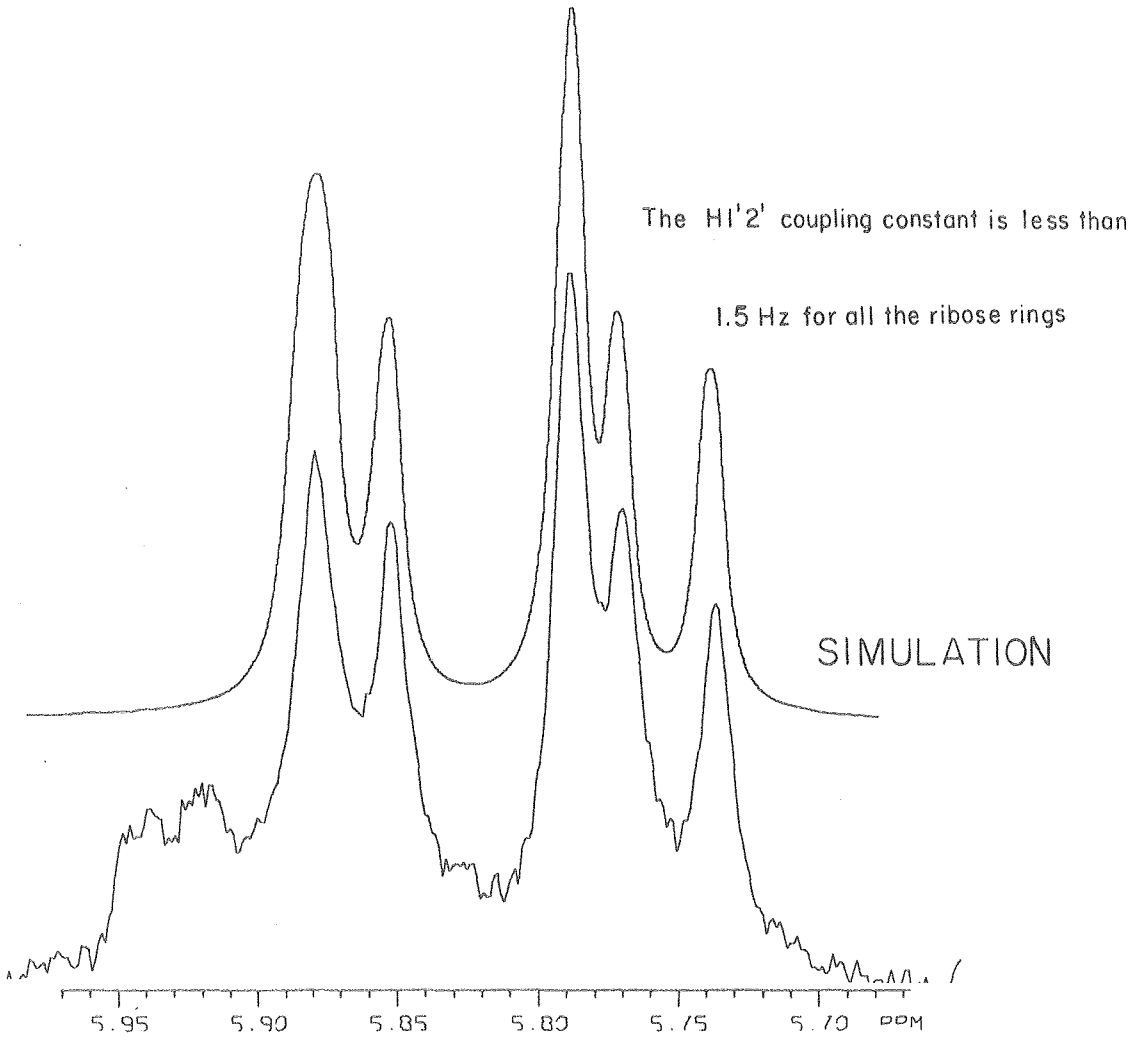
G) Linewidths of the Base Protons and Chemical Exchange

For the dCA_5G + dCT_5G system we see extensive broadening of some of the base protons during the melting transition. The fact that some of the resonances broaden while others stay very sharp is clearly seen in Figure 2.4. The reason for

Figure 2.21. Ribose H1' protons in dCT_5G + rCA_5G double helix at 5°C.

RIBOSE HI' PROTONS IN $dCT_5G + rCA_5G$ AT 5°C

All ribose sugars are in 3'endo conformation on the rCA_5G strand in the double helix



XBL 805-9520

this broadening is the fact that the system is undergoing chemical exchange between the double stranded helix (helix) and the single strands (coil). The rate of exchange, the population difference, as well as the difference in chemical shift between the two states are the factors which lead to broadening of the resonances.[45] Resonances such as the guanine (#1) H8 at 7.95 ppm, the thymines (#3-5) H6 at ∿ 7.6 ppm, all shift little, and therefore are sharp throughout the transition. The three interior adenine H2 protons (at 7.09, 7.08 and 7.67 ppm at 5°C) have such large chemical shift differences for the two states that their resonances are almost broadened into the baseline at 25°C.

The (#2) adenine H8 proton (8.32 ppm at 5°C) is well separated from its neighbors and broadens to moderate extent, which makes it[26,27] possible to extract rate constants for the helix-to-coil transition. Equation (1) applies when the system is just short of the fast exchange limit,[46] and has been used to calculate rate constants for the helix-to-coil transitions in other oligonucleotides.[26,27]

$$\frac{1}{\pi T_{2_{obs}}} = \frac{P_C}{\pi T_{2_C}} + \frac{P_H}{\pi T_{2_H}} + 4\pi P_C{}^2 P_H{}^2 (\nu_H - \nu_C)^2 (\tau_H + \tau_C) \qquad (1)$$

$1/\pi T_{2_H}$ and $1/\pi T_{2_C}$ are the measured linewidth of the helix and coil states, respectively, while $1/\pi T_{2_{obs}}$ is the observed linewidth in the broadened spectrum. The linewidth of the helix state for a particular proton is assumed to be

independent of temperature and is thus the measured linewidth of the helix at 5°C. The linewidth of the coil was taken from the measured linewidths during the melting of the single strand. ν_H and ν_C are chemical shifts of the protons in the helix and coil given in Hz, again ν_H is assumed to be the chemical shift of the helix at 5°C, while ν_C is measured from the melting of the single strand. P_H and P_C are the populations at a given temperature for the double and single strands, respectively. The populations were obtained from analysis of the absorbance versus temperature melting curve. τ_H and τ_C are the lifetimes of the helix and coil states; by the definitions, $P_H\tau_C = P_C\tau_H$.

We chose to calculate lifetimes of the helix and coil from the broadening of the #2 adenine H8 proton at 25°C, because at this temperature we can be fairly sure that equation (1) is applicable. Figure 2.22 shows that at 35°C the linewidth of this resonance is only 3.5 Hz, whereas for 25°C the linewidth is 16 Hz. At 35°C there are approximately equal populations of the helix and coil, so the system must be in fast exchange in order to get this linewidth. By 25°C the exchange lifetimes have changed to such an extent that one sees broadening of the peak and equation (1) can be used to model this broadening. At 25°C, from the data given in Table VIII, we calculate the lifetimes of the helix and coil to be 14 msec and 2 msec respectively. Similar studies

Figure 2.22. Linewidth versus temperature for #2 H8 adenine proton in dCA_5G for double and single strands.

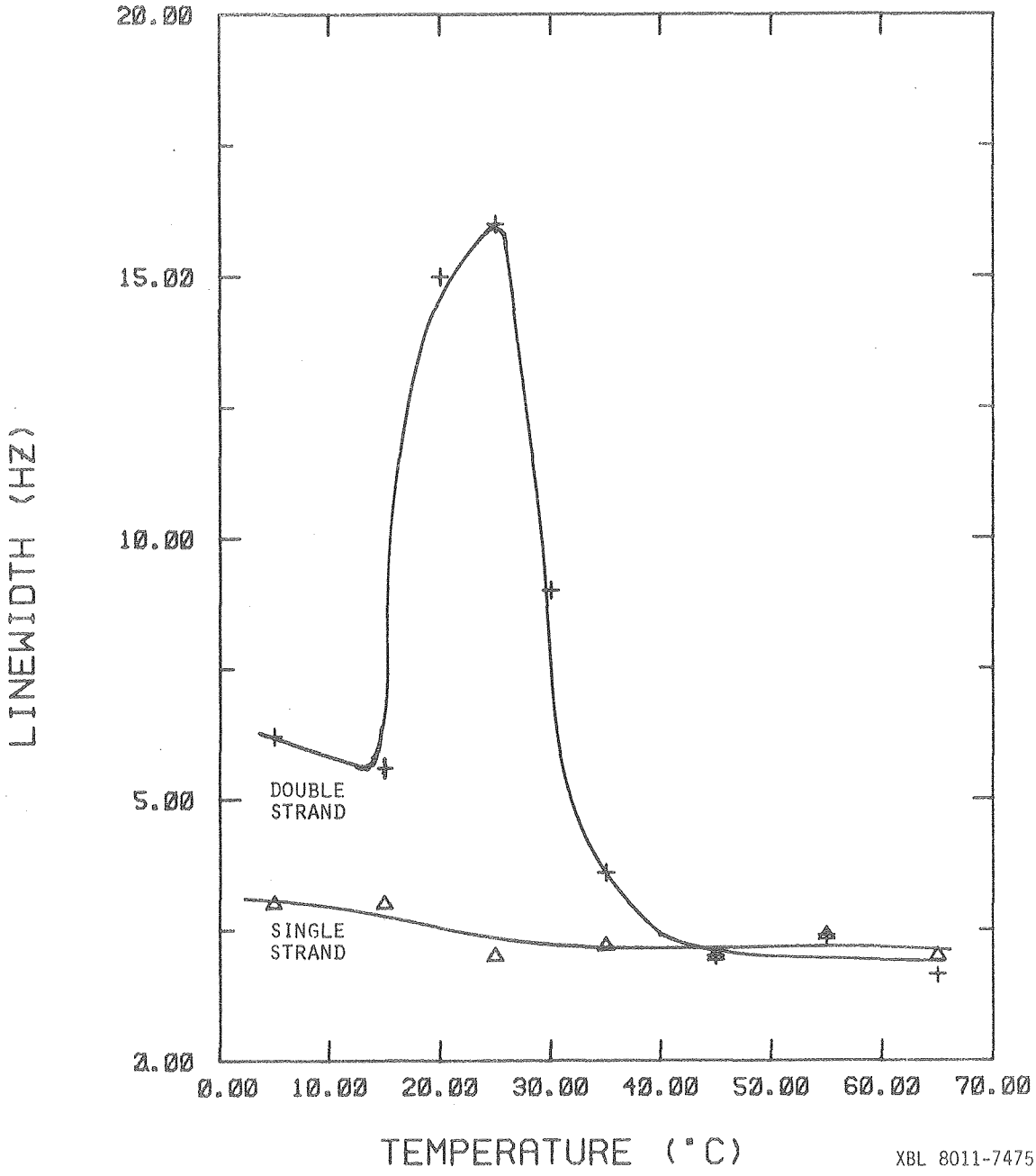

20.00
15.00
10.00
5.00
0.00
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
LINEWIDTH (HZ)
TEMPERATURE (°C)
DOUBLE STRAND
SINGLE STRAND
XBL 8011-7475

TABLE VIII

LINEWIDTH DATA FOR #2 ADENINE H8 PROTON

IN $dCA_5G + dCT_5G$ DOUBLE HELIX

360 MHz at 25°C

$1/\pi T_{2obs} = 15$ Hz $1/\pi T_{2c} = 2$Hz $1/\pi T_{2H} = 6$ Hz

$p_H = 0.88$ $p_c = 0.12$ $\nu_H - \nu_c = 72$ Hz

Using equation (1) and $p_H\tau_c = p_c\tau_H$ we calculate

$\tau_H = 14$ msec and $\tau_c = 2$ msec

270 MHz at 24°C

$1/\pi T_{2obs} = 15$ Hz $1/\pi T_{2c} = 2$Hz $1/\pi T_{2h} \approx 6$ Hz

$p_H \cong 0.88$ $p_c \cong 0.12$ $\nu_H - \nu_c = 55$ Hz

we then calculate

$\tau_H = 19$ msec and $\tau_c = 2.5$ msec

were done on the 270 MHz instrument at 24°C with the data given in Table VIII. The lifetimes are probably not very meaningful at temperatures other than 25°C because at higher temperatures the broadening caused by exchange is not much larger than the natural linewidth of the peak, so the errors involved in calculating the lifetimes are very large. At temperatures lower than 25°C the assumption of close to fast exchange and the use of equation (1) may not be valid. The fact that the lifetimes calculated at 25°C from data obtained at two different frequencies are in good agreement, is indicative that at this temperature equation (1) is applicable to our system. These lifetimes and their rate constants are consistent with kinetic results obtained on other oligonucleotides by temperature jump methods.[48,49]

H) Base Pair Overlap in the Double Strand

The chemical shifts of the nonexchangeable protons for the helix $dCA_5G + dCT_5G$ have been compared with the calculated chemical shifts assuming a B form DNA geometry. The calculations for the double helix were done using the model by Arter and Schmidt,[41] which incorporates next nearest neighbor effects as well as the effects of the closest base pairs. The data shown in Table VI gives the chemical shifts for a B DNA geometry and the observed shifts for this helix. The prediction of the chemical shifts of the base protons in A RNA and A'RNA geometry using Arter and Schmidt's

numbers are also given in Table VI.[41] One can see that these predictions do not give a better fit to the experimental numbers than the B DNA form.

The observed chemical shifts of the base protons which we were able to assign for the helix $dCT_5G + rCA_5G$ at 5°C are shown in Table VI. These can then be compared with predicted chemical shifts from the three geometries also given in Table VI. None of the fits between the calculated and the experimental data are very good for the base protons, so the imino protons seem to be more useful in these comparisons.

4. DISCUSSION

A) Comparison of the Structure of Oligonucleotide Complexes In Solution

I. The double Stranded Structures

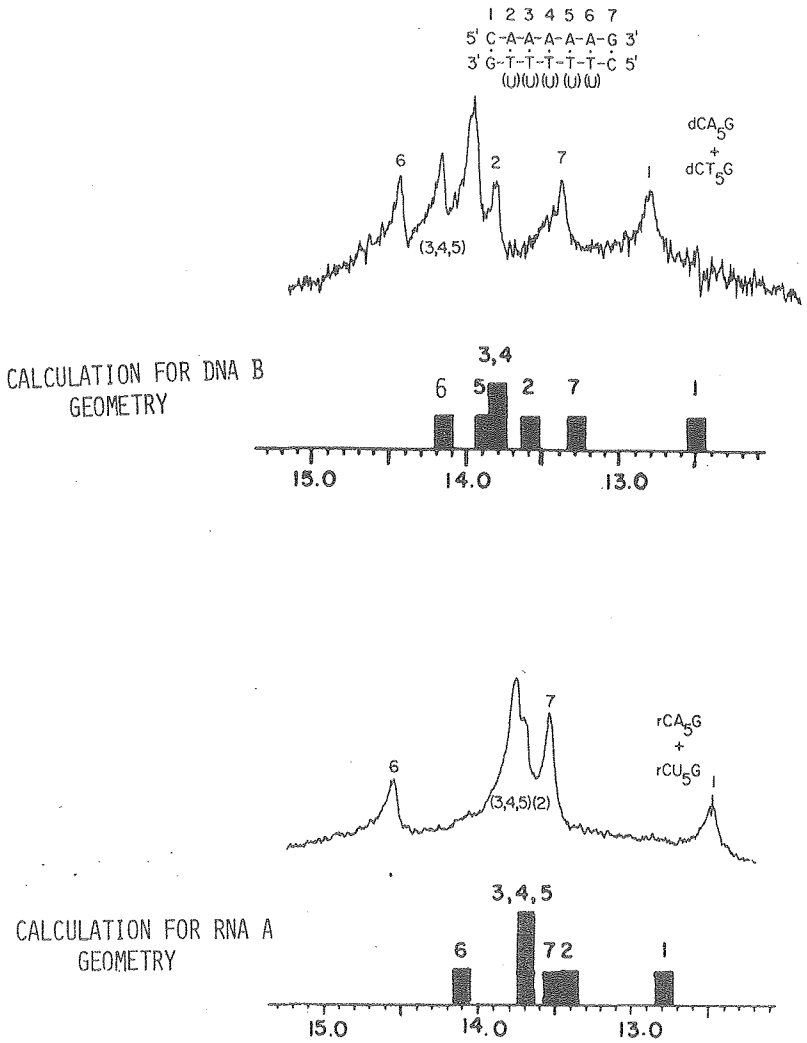
Figures 2.7a-c show the low field spectra of the three double stranded systems at 5°C. There are seven imino protons in each spectrum, one for each base pair in the double helix. From examining the chemical shifts of the imino protons in the three double strand, one sees distinct differences in the peak positions of many protons. This is indicative of different conformations for each system. The best fits for calculated and experimental results are for the $dCT_5G + dCA_5G$ in a DNA B type geometry, while the $dCT_5G + rCA_5G$ and $rCU_5G + rCA_5G$ are in an RNA A or A' type geometry.

The fits of the DNA B form to the deoxyribose double helix is excellent in the relative positions of the imino protons, but are not as good in their predictions of the exact positions, as seen in Table VII and Figure 2.23. The predicted resonances are all about 0.3 ppm higher field than the experimental results. Since the conformations of RNA and DNA helices are different, it is likely that the intrinsic positions of these isolated base pairs are also different. As more proton NMR work is done on deoxyribo-oligonucleotides, it may be possible to obtain empirical results for the chemical shifts of the isolated A·T and C·G imino protons in deoxyribose systems.

RNA structures are found to be more rigid and less susceptible to conformational changes than their DNA counterparts.[3] We would then expect that the RNA helix in our study would be close to an RNA A form geometry. As seen in Table VII and Figure 2.23, there are large differences between the experimental results and the calculations of the chemical shifts for an RNA A or A' geometry. These differences can be due to several factors such as inaccurate values for the ring currents used in the calculations, incorrect geometries assumed in the calculations, sequence dependent geometries, or other parameters besides ring currents having significant effects on the chemical shifts of the imino protons (see Borer et al., 1975, for discussion of these effects).[15] Aggregation will also directly effect the chemical shifts of the terminal base

Figure 2.23. Comparison of experimental and calculated chemical shifts of the imino protons.

COMPARISON OF CALCULATED AND EXPERIMENTAL CHEMICAL SHIFTS OF THE IMINO PROTONS



XBL 8011-7491

pair protons. Equilibrium sedimentation on similar oligonucleotides have shown that there is significant aggregation at these concentrations, so the chemical shift of the terminal base pairs may be affected.[44]

The chemical shifts of the base protons are compared with the calculated results for several of the oligonucleotide double helices in Table VI. One sees large discrepancies between the experimental and calculated chemical shifts for the $dCA_5G + dCT_5G$ system. We are investigating the discrepancies to see if the differences are due to the oligonucleotide being in a slightly different geometry than the classical DNA B form. Calculations of the chemical shifts for other energy minimized geometries are presently being performed (D. Keller, unpublished) to see how slight changes in helical parameters, such as winding angle, twist, and tilt, effect the chemical shifts of the base protons.

The coupling constants and chemical shifts of the ribose and deoxyribose protons have been extensively used in the past to obtain conformational information on oligonucleotides in solution.[8,9] One of the most meaningful parameters in the geometry of the nucleic acid is the conformation of the sugar ring.[50] In the $dCT_5G + dCA_5G$ double helix, two of the deoxyribose sugar rings in the dCA_5G strand were found to be 90% in a 2' endo conformation. The 2' endo sugar puckers are found in B form geometries which is consistent with this helix being in a B type conformation.

The hybrid dCT_5G + dCA_5G was also well enough resolved in the H1' proton region to obtain coupling constant data. The evidence of a 3' endo conformation of the ribose strand is indicative of an A type geometry. Knowing the sugar pucker of the helix greatly restricts the range of conformations for the sugar phosphate backbone, thereby ruling out many possible geometries for the helix. This fact is especially useful in conjunction with the ring current calculations, and may make it possible to define conformations more exactly than the qualitative "A" type or "B" type geometries.

II. Triple Strand

The system rCU_5G + dCA_5G forms a triple strand under our conditions. The low field imino proton spectrum of this mixture is quite different than the other spectra shown in Figure 2.7. For example, there is a broad resonance at very low field around 15.0 ppm in Figure 2.7d, the normal Watson-Crick imino protons are not found to resonate higher than $\sim$ 14.6 ppm. This resonance as well as those centered at 14.6 ppm are most likely due to reverse Hoogsteen base pairs involved in the triple strand.

Geerdes and Hilbers have studied [oligo A - $(\text{oligo U})_2$] triplexes in solution by ^{1}H NMR.[23] From the imino protons in this system they concluded one assignment for the intrinsic position of the Watson Crick A·U pair in the triple strand would be $\sim$ 14.3 ppm, whereas the Hoogsteen pair was found

at 14.8 ppm. Robillard and Reid have recently attempted calculations which empirically optimize the magnitudes of the chemical shifts of the isolated base pairs by a method which assumes a strong similarity between the crystal and solution structure of tRNA.[6] They found the optimized intrinsic position of the reversed Hoogsteen A·U to be at 14.9 ppm. Kallenbach et al. have studied triple stranded structures in which oligo U_{15} and rAMP formed two sets of imino resonances, presumably one for the Watson-Crick and one for the reverse Hoogsteen base pairs.[19] They concluded that the A·U reverse Hoogsteen base pair's most probable intrinsic chemical shift was 14.1, but were unable to rule out an alternate assignment which would have switched the position of the two types of base pairs, and given the intrinsic shift of the reverse Hoogsteen as $\sim$ 14.8 ppm.

In our system we are unable to make unambiguous assignments of the imino protons, but we think that the broad resonance at $\sim$ 15.0 ppm is due to an A·U reverse Hoogsteen base pair. It seems to resonate at too low field to be from a normal Watson-Crick type A·U resonance, and the above studies on triplexes indicate that the reverse Hoogsteen base pair's intrinsic shifts may be in the region of 15.0 ppm. This would indicate that the other resonances at 14.2 - 14.5 ppm would also be mostly from reverse Hoogsteen base pairs. The intrinsic shift of an rU·dA·rU reverse Hoogsteen pair would then be $\geq$ 15.0 ppm.

The fact that there are so many types of resonances in the low field region of the $dCA_5G + rCU_5G$ system, many more than the seven protons one sees in the double helical systems, indicates that the oligomers are fully or partially in the triple strand conformation.

B) Analysis of Thermodynamic Parameters by Chemical Shift Versus Temperature Plots of the Base Protons

The validity of the T_m or enthalpy values from melting curves is often dependent upon the method of analysis. In order to get useful results from a melting curve, one must estimate the temperature dependence of the double and single stranded molecules.[15,17] One major advantage in using non-selfcomplementary oligonucleotides is that the properties of the single strand can be studied separately from the double to single strand transition. The temperature dependence of the chemical shifts of the base protons in single stranded molecules have been studied in many systems. In some cases the changes in chemical shifts of the base protons are linear with temperature,[16] but for many other protons there are large deviations from linearity (see Figures 2.12 - 2.14 and Shum).[51] For non-selfcomplementary molecules, one can experimentally obtain the temperature dependence of the chemical shifts of the single strands, while for selfcomplementary molecules, one is forced to assume linearity for the chemical shifts. This assumption can lead

to large errors in the T_m. For example in the system $dCA_5G + dCT_5G$ the analysis of the curves with and without the experimental temperature dependence of the single strand gives T_m values which differ by up to 5°C for some of the protons. This problem seems to be much more important in NMR melts than optical studies, and should be taken into account when interpreting thermodynamic parameters derived from NMR melting curves.

There is a similar problem involved in the estimation of the lower baseline of the melting curve. The temperature dependence of the chemical shifts of the double strand may not be linear with temperature, and this can also lead to errors in the T_m.

Difficulties involved in obtaining equilibrium constants from chemical shift data have recently been discussed.[52] One usually assumes that the system is always in fast exchange, on the NMR timescale, and that the measured chemical shift is a weight average of all the states. The validity of this analysis breaks down when the system is not in the fast exchange domain, which happens during the melting of most oligonucleotides. The errors made in assuming fast exchange can be large in many cases.

We have constructed a model system which calculates the lineshape of the resonances for any exchange rates linking two states. The expression used is essentially that derived

by Gutowsky and Holm,[45] except that their assumption of the linewidth of the two states being equal was not made. The expression for the intensity of absorption for a two state system with exchange is given by

$$\nu(f) = \frac{-K[J(1 + \tau(p_A\omega_B + p_B\omega_A)) + QR]}{J^2 + R^2} \qquad (2)$$

where

$$\tau = \tau_a\tau_b/(\tau_a+\tau_b) \qquad \omega_n = \pi(\text{linewidth of state n})$$

and

$$J = \tau[\omega_A\omega_B - (\pi(f_A+f_B) - f)^2 + \pi^2(f_A-f_B)^2] + p_A\omega_A + p_B\omega_B$$

$$Q = \tau[\pi(f_A+f_B) - f - \pi(p_Ap_B)(f_A-f_B)]$$

$$R = \pi(f_A-f_B)(p_A-p_B) + \tau\pi(f_A-f_B)(\omega_B-\omega_A) + [\pi(f_A+f_B) - f][1 + \tau(\omega_A+\omega_B)]$$

where τ_n, f_n, and p_n are the lifetime, the frequency (in Hz), and population, respectively, of state n.

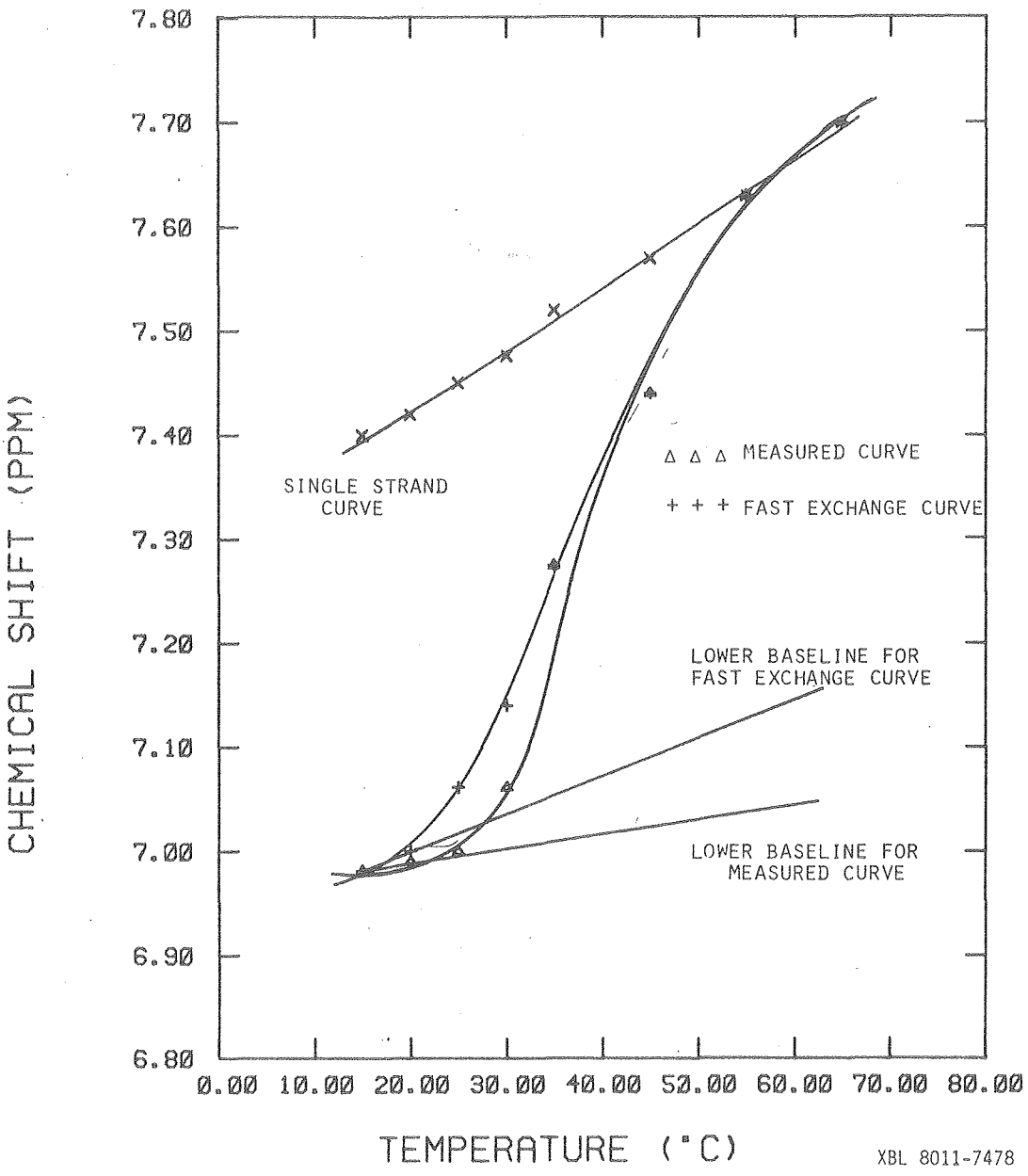
The program which calculates the lineshape of a two state system with exchange is called WIDTH2 and given in Appendix II. The program uses equation (2) plotted with a program called PLOT written, and kindly provided by Mr. David Keller. WIDTH2 also finds the chemical shift and linewidth of the largest peak in the system. These can then be compared with the experimental results, and the parameters such as exchange rates varied until a good fit is found between the experimental and calculated results.

The program WIDTH2 was modified to a program called ITWIDT which changes the mean exchange lifetime τ (where $\tau = \tau_a\tau_b/(\tau_a+\tau_b)$ until the best fit is found between the calculated and experimental chemical shift of the largest peak. For a given population of the two states it iterates until the closest fit is found between the experimental and calculated results.

These programs were used to model an NMR melting curve to see how large the errors can be in extracting thermodynamic data by always assuming fast exchange. If the system were in fast exchange throughout the whole helix-to-coil transition, then the chemical shift observed is a weight average of the two states. The observed chemical shift would then be a direct measure of the populations of the two states (assuming one knows the chemical shift of the helix and the coil). For this analysis we will refer to the measured chemical shifts as those which are found if the system were from equation 2. If the fast exchange chemical shift and measured chemical shift are different, then for that population and exchange rates the normal analysis of the melting curve is not valid.

Figure 2.24 shows the melting curves for a model system which mimics the behavior of an interior adenine H2 proton in $dCA_5G + dCT_5G$. What was found was that the fast exchange chemical shifts agree well with the measured chemical shifts

Figure 2.24. Comparison of melting curves when system is in fast exchange and when it is no longer in the fast exchange region (measured curve).


7.80
7.70
7.60
7.50
7.40
7.30
7.20
7.10
7.00
6.90
6.80
CHEMICAL SHIFT (PPM)
0.00 10.00 20.00 30.00 40.00 50.00 60.00 70.00 80.00
TEMPERATURE (°C)
SINGLE STRAND CURVE
Δ Δ Δ MEASURED CURVE
+ + + FAST EXCHANGE CURVE
LOWER BASELINE FOR FAST EXCHANGE CURVE
LOWER BASELINE FOR MEASURED CURVE
XBL 8011-7478

ΔΔΔΔΔ

above the T_m of the system. This is because the system actually is in fast exchange over this temperature range for all the protons. Below the T_m, many of the protons are no longer in fast exchange, so we see lines which are broadened due to exchange. The chemical shift is no longer the weight average, but skewed toward the chemical shift of the state with the largest population. This will have the effect of raising the T_m above its real value. In most of the model systems studied this effect was fairly small, with the largest change being 2 - 4°C in the T_m. The error is in the wrong direction to explain the 5°C discrepency between the optical and the NMR T_m of the $dCA_5G + dCT_5G$ system, and this effect does not seem to be important in our case.

Another error produced in assuming fast exchange which is probably much larger in our system involves the estimation ofthe lower baseline in the melting curve. When the system is 80-100% in the double strand, many of the protons are near the slow exchange limit so there should be two peaks observed, weighted according to the populations of the two states. It will often be difficult, if not impossible, to see the peak which corresponds to a state with less than 20% population, due to the large number of resonances and the fact that many of them may still be exchange broadened. Thus the chemical shift that we measure will only be the chemical shift of the state with population of > 80%, and not a weight average

chemical shift. This error will affect the shape of the lower baseline drawn in the analysis of the melting curve. The measured curve will overestimate the population of the double helix and thus make the lower baseline slope less than it actually is. As seen in Figure 24, this has the effect of lowering the T_m of the measured curve below the real T_m.

In the system dCA_5G + dCT_5G, the T_m found optically had a value 5°C higher than the T_m obtained by NMR, which showed all the resonances melting at 28-30°C. The reason for this difference is the invalid assumption of fast exchange throughout the transition, which causes an incorrect extimation of the lower baseline of the NMR melting curve. In these molecules we took all the lower baselines to be flat, so that the error is probably systematic and will be similar for all the protons. Modeling this system gives the same conclusion, that the errors due to assuming fast exchange are approximately the same for all the protons. Therefore we think that for dCT_5G + dCA_5G the terminal and interior base pairs melt at the same temperature, within 2 - 3°C, and that an all-or-none transition is a good approximation for this system.

C) Fraying Versus Melting of Oligonucleotides

It is important to understand the distinction between fraying of the ends and the differential melting of the ends of an oligonucleotide. Here we refer to the melting of a base pair in the oligomer with respect to the fraction of the base pair which is formed, or involved in hydrogen bonding. An 80% melted base pair would have an equilibrium concentration of oligonucleotides with 80% of the base pair broken and 20% of the base pair formed. Melting thus reflects an equilibrium effect and is dependent only upon the concentration of the two states.

Fraying is defined as the rapid opening and closing of a base pair.[53-55] It is thus a kinetic effect and the important parameters involved are the rate constants linking the open and closed states. Fraying can manifest itself in the H-bonding imino protons where the exchange rate of the proton is reflected by the linewidth of the resonance (see Crothers et al.,[56] and Hilberts,[53] for discussion of exchange of imino protons with H_2O). In our oligomer systems the C-G base-paired imino protons broaden and disappear before the interior base pairs, which is indicative of fraying of the ends of the helix. This does not mean the end base pairs are melted at the temperature where the resonance has disappeared. For example, in the helix $dCT_5G + dCA_5G$ at 25°C, the imino proton region is extremely broad, and by 30°C the resonances have disappeared into the

baseline (see Figure 2.9). This is not representative of the equilibrium concentrations of these states, the oligomer is fraying at this point but it is not melted. In fact, the chemical shift versus temperature data of protons on all bases and optical studies of this oligomer show that it is > 50% in the double strand at 30°C, where all the imino protons have disappeared. The temperature dependence of the broadening of the imino protons ismainly a kinetic, not an equilibrium, effect. Thus an end base pair can be fraying and not "melted" at the same time. In our system, although we see fraying of the ends of the helix, i.e., the end bases are opening and closing faster than the interior base pairs, the ends do not melt appreciably lower than the rest of the helix. It is worth noting that fraying and melting reflect not only different processes, but there may also be different states involved in the two processes. For example, in order for an imino proton to exchange with water, the base pair must open to some extent. This opened state most likely differs from the "melted" state in a molecule with differential melting on the ends of the helix. Care should be taken in distinguishing between fraying and melting, as well as the effects ascribed to them.

5. CONCLUSION

We have studied the following oligonucleotides separately, and in their complementary mixtures by proton NMR: dCA_5G, dCT_5G, rCA_5G and rCU_5G. Results on the single strands show that the ribose and deoxyribose strands of CA_5G have different conformations in solution, and specifically that there is more base-base stacking in the deoxyribose strand.

Results on the mixtures indicate that rCU_5G + dCA_5G at least partially forms a triple strand with a 2:1 rCU_5G:dCA_5G helix, while rCU_5G + rCA_5G, rCA_5G + dCT_5G, and dCT_5G + dCA_5G form double helical structures. Comparison of the three double strands shows that they all have different structures in solution. The chemical shifts of the imino protons and the calculations of sugar ring pucker in these systems indicate that the deoxyribose helix (dCA_5G + dCT_5G) is in a B type geometry, whereas the ribose (rCA_5G + rCU_5G) and hybrid (rCA_5G + dCT_5G) helices are in an A or A' type geometry.

The melting of the helix dCA_5G + dCT_5G as followed by the chemical shift changes of the nonexchangeable base protons shows that base pairs on the ends of the helix melt at approximately the same temperature as the interior base pairs. The temperature dependence of the imino protons indicates there is fraying of the ends of the helix. Thus this helix exhibits fraying of the ends, but very low concentrations of partially formed helices at the same time. These results

indicate that a two-state model is a good approximation of the helix-to-coil transition in this oligonucleotide, as often assumed in optical studies. Some of the biological implications due to the differences in stability of these oligonucleotides have been discussed elsewhere.[24] The differing conformations for these DNA, RNA, and hybrid helices are probably also important in enzymatic recognition of different types of helices.

Studies on the kinetics of these oligonucleotide helices, as well as the changes in the conformation and stability of double helical oligonucleotides when there is a mismatched base on one strand, or when frameshift mutagens are bound, are presently being pursued.

CHAPTER III

1. INTRODUCTION

It is important to study the exchange of the H-bonding protons in nucleic acids if one is to understand their stability, structure, and flexibility in solution. In order for a proton to exchange, the H-bond must be broken which will require at least some change in the conformation of the helix. The breathing of a helix and the structural stability of a particular region may be important in enzyme-nucleic acid interactions, and also mutational events.[39] Disruption of the base pairs will be important in transcription and replication of DNA.

Exchange in polynucleotides has been studied by tritium exchange experiments,[57-59] and also by H-D exchange as monitored by a kinetic difference spectrum in the UV. These results have been recently reviewed by Kallenbach et al.,[60] and demonstrate the importance of the dynamics and fluctuations of DNA and RNA.

Proton NMR has also been used to study exchange of H-bonded protons in tRNA and oligonucleotides. (See Hilbers[53] for a review of exchange in nucleic acids as studied by NMR.) Crothers et al.[56] studied the thermal unfolding of $tRNA^{fMet}$ by the broadening of the low field imino protons in the molecule. The broadening of these resonances is due to exchange with water, and the equations used to relate the width and chemical shift of the resonance to exchange rates are given there.

Studies of exchange in oligonucleotides have been done by Patel and Hilbers on dATGCAT,[54,55] and by Kallenbach et al.[19] on dAAAGCTTT, by monitoring the behavior of the imino protons. They concluded that the internal G imino proton exchanged in an open limited case, which means the rate of exchange is limited by the opening of the base pair and not by exchange with H_2O. For the other A·T base paired imino protons, exchange took place from partially formed double helical states. This indicates fraying of the ends of the helix, which means the ends of the helix were rapidly opening and closing.

These earlier studies of exchange by NMR involved interpreting the linewidth and chemical shift of the imino protons. Redfield developed the method of transfer of saturation to study exchange rates of H-bonded protons with H_2O.[61-63] This method involves perturbing part of the system, and then monitoring the system as it exchanges with an unperturbed part of the system. Johnston and Redfield studied exchange rates in tRNA by this method.[62,63]

In this chapter we will investigate the exchange of the imino protons on a deoxyribo-oligonucleotide $dCA_5G + dCT_5G$, and also a hybrid helix $dCT_5G + rCA_5G$. The fraying on the ends of the helix as well as the exchange of protons on the interior of the helix will be discussed.

2. Experimental Methods

The samples used in these studies were the same ones used to study the imino protons in the oligonucleotides. The experiments were done on the 360 MHz instrument at Stanford Magnetic Resonance Laboratory, and the 270 MHz instrument at the Laboratory of Chemical Biodynamics. The Redfield 214 pulse sequence setup and the saturation recovery experiment are routinely done on the Stanford facility, and are well documented there. For the 270 MHz instrument slight modifications of the existing setup had to be made in order to do the Redfield pulse sequence and the saturation recovery experiments. Since this procedure is not documented elsewhere,[64] the modifications made on the system will be discussed later.

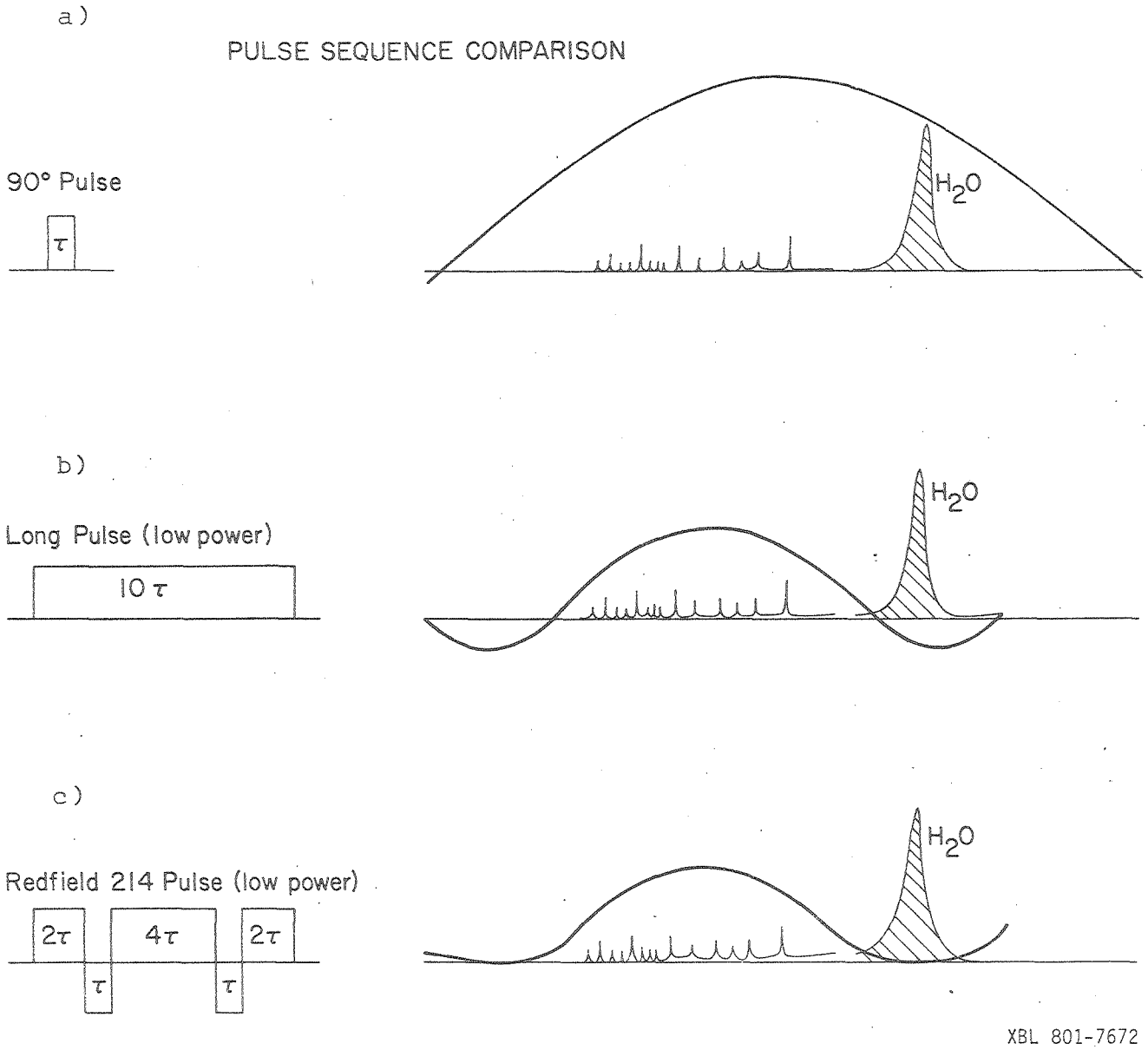
A description of the theory of the Redfield 214 pulse sequence is given elsewhere.[65] The idea of the pulse sequence stems from the fact that the frequency width of a radio frequency (Rf) pulse is inversely proportional to the length of the pulse. A typical 90° pulse is usually on the order of 2 - 30 μsec, and covers a much wider frequency range than the normal absorption frequencies of protons in solution. If one wants to reduce the frequency width of the Rf pulse, then the time of the pulse is simply increased. One method of water suppression similar to the Redfield sequence is a long low power pulse. The Rf pulse

Figure 3.1. Comparison of the shapes of an Rf pulse for:

a) a 90° pulse.

b) a long low power pulse.

c) a Redfield 214 pulse.


a)
PULSE SEQUENCE COMPARISON
90° Pulse
τ
H2O
b)
Long Pulse (low power)
10 τ
H2O
c)
Redfield 214 Pulse (low power)
2τ
τ
4τ
τ
2τ
H2O
XBL 801-7672

is centered at the frequency of the signals of interest, (they must be > 1 ppm away from the water signal), and the length of the pulse is adjusted so that the null in the Rf pulse is on top of the water signal. This reduces the intensity of the H_2O signal relative to the rest of the spectrum. The shape and the frequency width of a short 90° pulse and a long pulse are schematically shown in Figures 3.1a and 3.1b. Redfield improved the suppression of the water signal by broadening the null in the Rf pulse with use of two 180° phase shifts during the pulse. The shape of the Redfield 214 pulse is compared schematically with the long pulse in Figure 3.1.

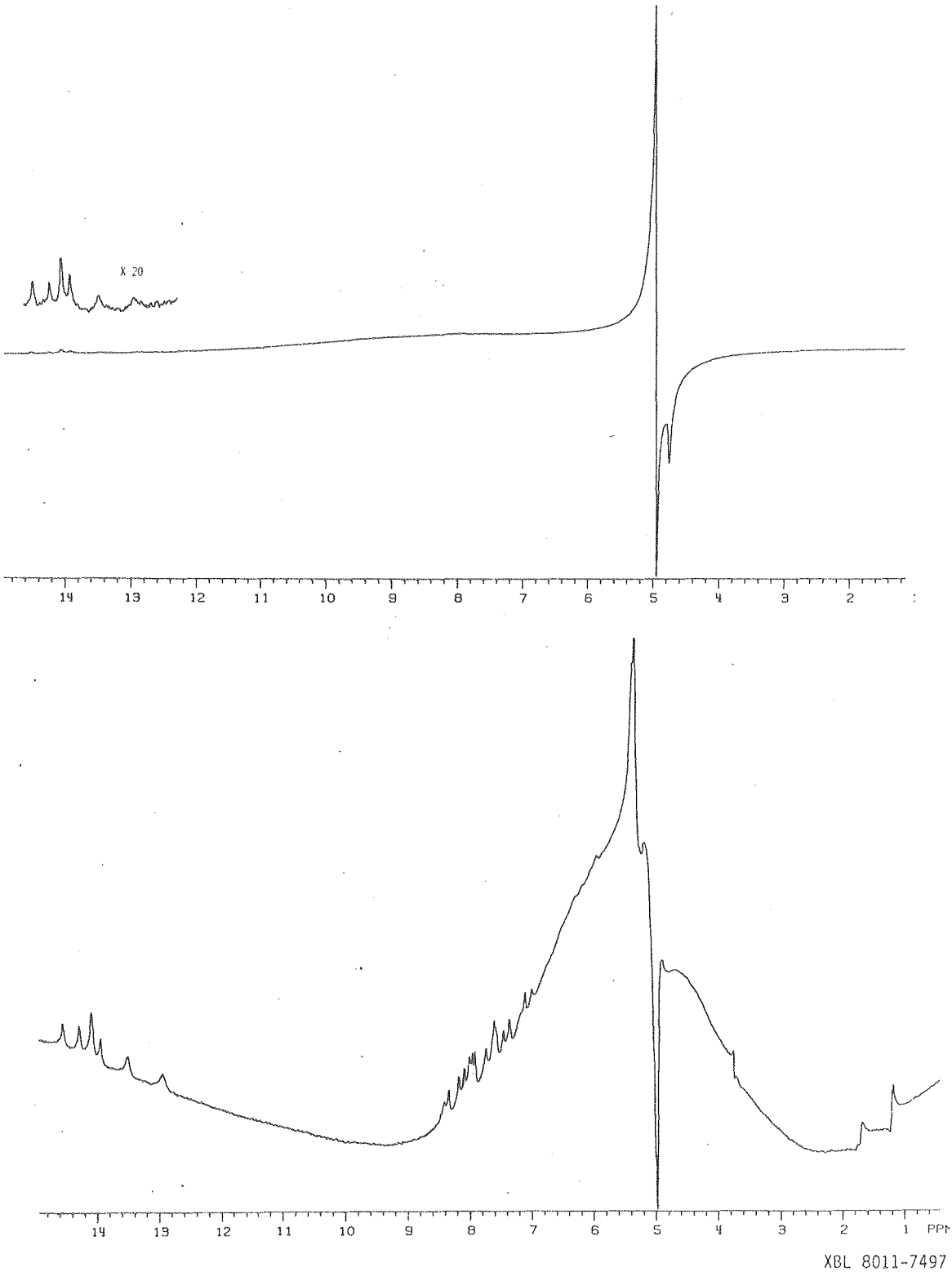
The Redfield sequence allows the H_2O signal to be reduced approximately 100 fold in most cases, thus cutting down the huge dynamic range problem encountered when doing ^1H NMR in H_2O. Audio filtering can also be used to further reduce the H_2O signal, and in most of our saturation recovery experiments the cutoff filter was set so as to filter the water signal, reducing its magnitude. The problem with this procedure is that it causes baseline distortions and phasing problems, over the area being filtered. Since the imino protons that we are interested in are 8-9 ppm downfield from water, there was little pertubation in the imino proton region, but it is impossible to correctly phase the whole spectrum. Another method of audio filtering which is far

superior to the cutoff filter in these experiments, is the use of a tunable notch filter to suppress the H_2O signal.[67] The notch filter reduced the water signal to a much larger extent, and also induced far less baseline distortion than the cutoff filter. Most of the saturation recovery experiments described here were done using the cutoff filter since the notch filter was not yet available. Comparison of the size of the water signal in the Redfield sequence with and without the notch filter is shown in Figure 3.2.

For the 270 MHz NMR spectrometer, instrumental details for normal Fourier transform operation are given elsewhere,[64] and only modifications needed for these experiments will be discussed. The Redfield 214 pulse sequence requires a long low power pulse, as well as the ability to make two 180° phase shifts during the pulse. In our experiments a low power linear amplifier was used for the Redfield 214 pulse in place of the normal high power proton amplifier. The same amplifier was also used in the saturation recovery experiments on the decoupling channel to saturate the imino resonance. Figure 3.3 shows the experimental setup for the observe channel when using the Redfield pulse sequence. Figure 3.4 shows the setup for the decoupling channel when running the saturation recovery experiments.

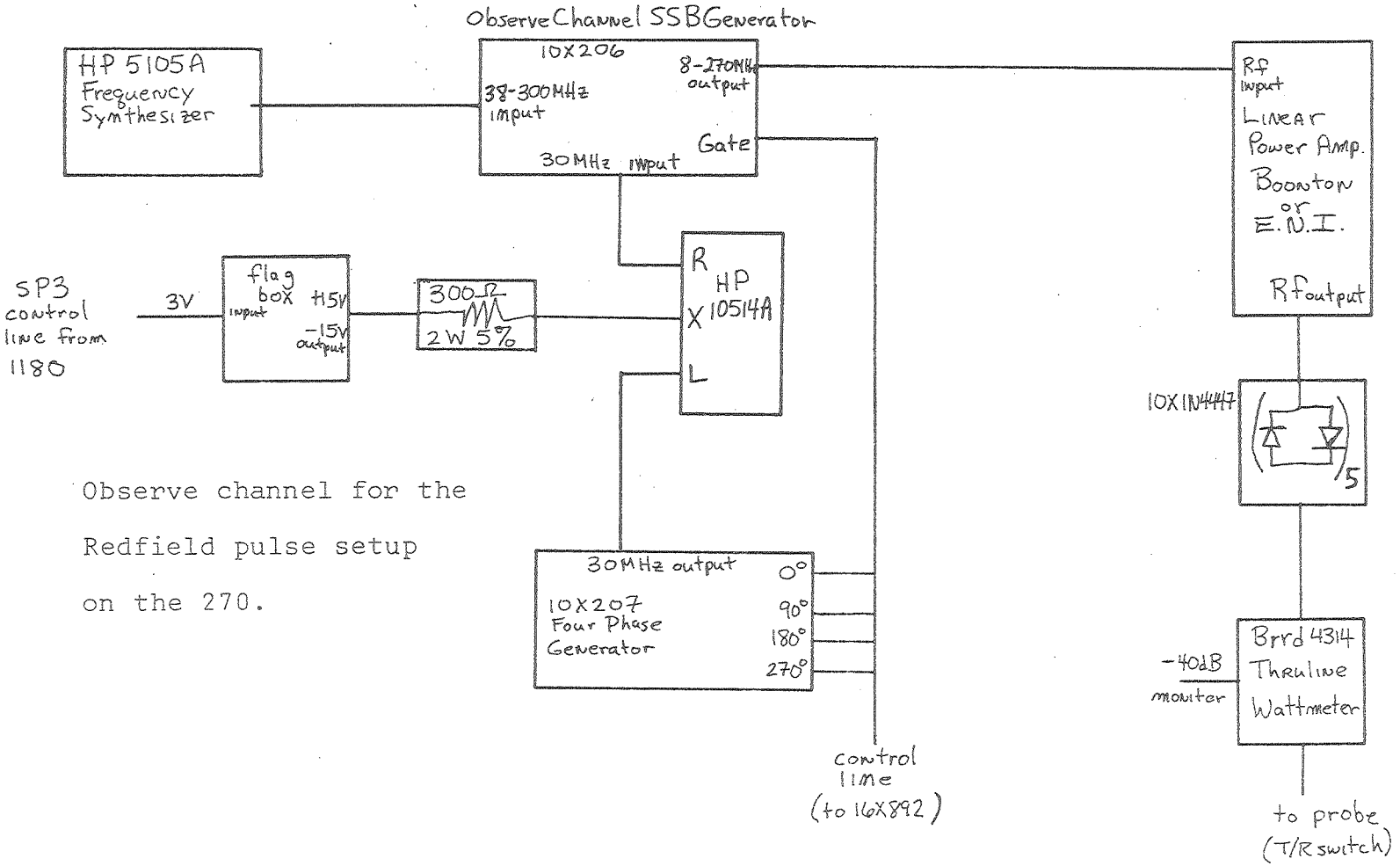
When using the Redfield sequence, all the protons over a frequency range are not equally excited. This can present problems in comparison of the relative intensities of peaks

Figure 3.2. Comparison of Redfield sequence spectra without notch filter (top) and with notch filter (bottom).



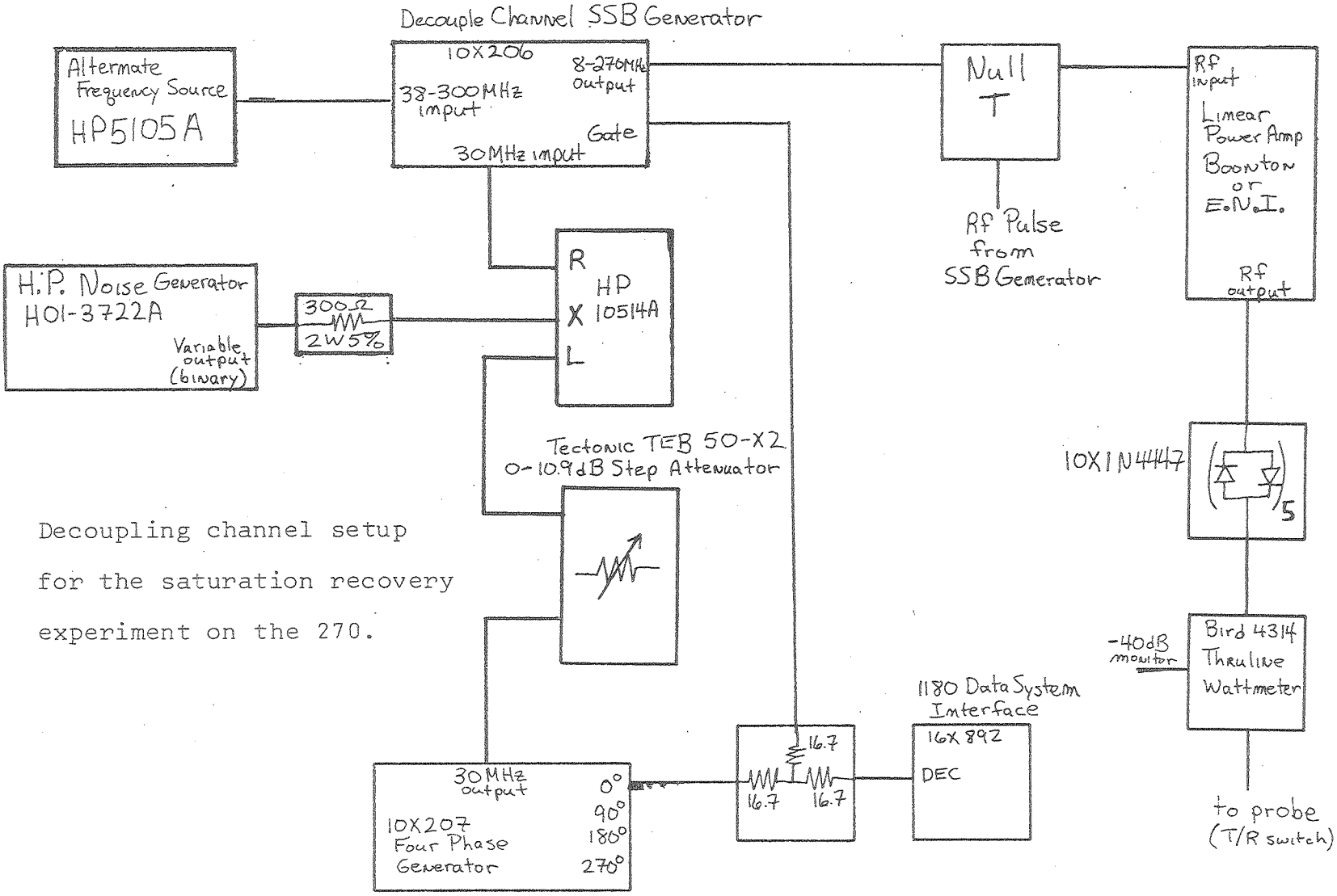
XBL 8011-7497

Figure 3.3. Block diagram of the observe channel for the Redfield pulse sequence setup on the 270 MHz NMR spectrometer.



Observe channel for the Redfield pulse setup on the 270.

Figure 3.4. Block diagram for the decoupling channel in the saturation recovery experiment on the 270 MHz spectrometer.



Decoupling channel setup for the saturation recovery experiment on the 270.

in different parts of the spectrum. For this reason the center of the Rf pulse was placed as close to the center of the imino proton region as possible. Over the three ppm range of the imino protons in nucleic acids, there is only a small intensity difference in the Rf pulse, typically less than 5%, so there was little effect on the intensities of the peaks due to the pulse width. Figure 2.8 shows the integration of the imino proton region of the DNA helix.

In order to find optimum conditions for the Redfield pulse sequence, th power of the Rf pulse must be changed. A 90° pulse is normally obtained by changing the length of the pulse and holding the power of the pulse constant. Equivalently, one can keep the pulse length constant and change the power. The Redfield sequence requires a pulse length which is set by the frequency difference between the center of the pulse and the H_2O signal, so one must change the power of the pulse in order to change the flip angle. The optimum power for the Redfield pulse of a given length was found by maximizing the signal to noise of the 3 imino proton in a sample of 50 mM 3' UMP at pH = 4.5 in H_2O. This signal is found around 10.8 ppm at this pH and 25°C. The power level was increased until the largest signal to noise was found; whether this power level represents a 90° pulse under these conditions is not known, but higher power levels gave lower signal to noise.

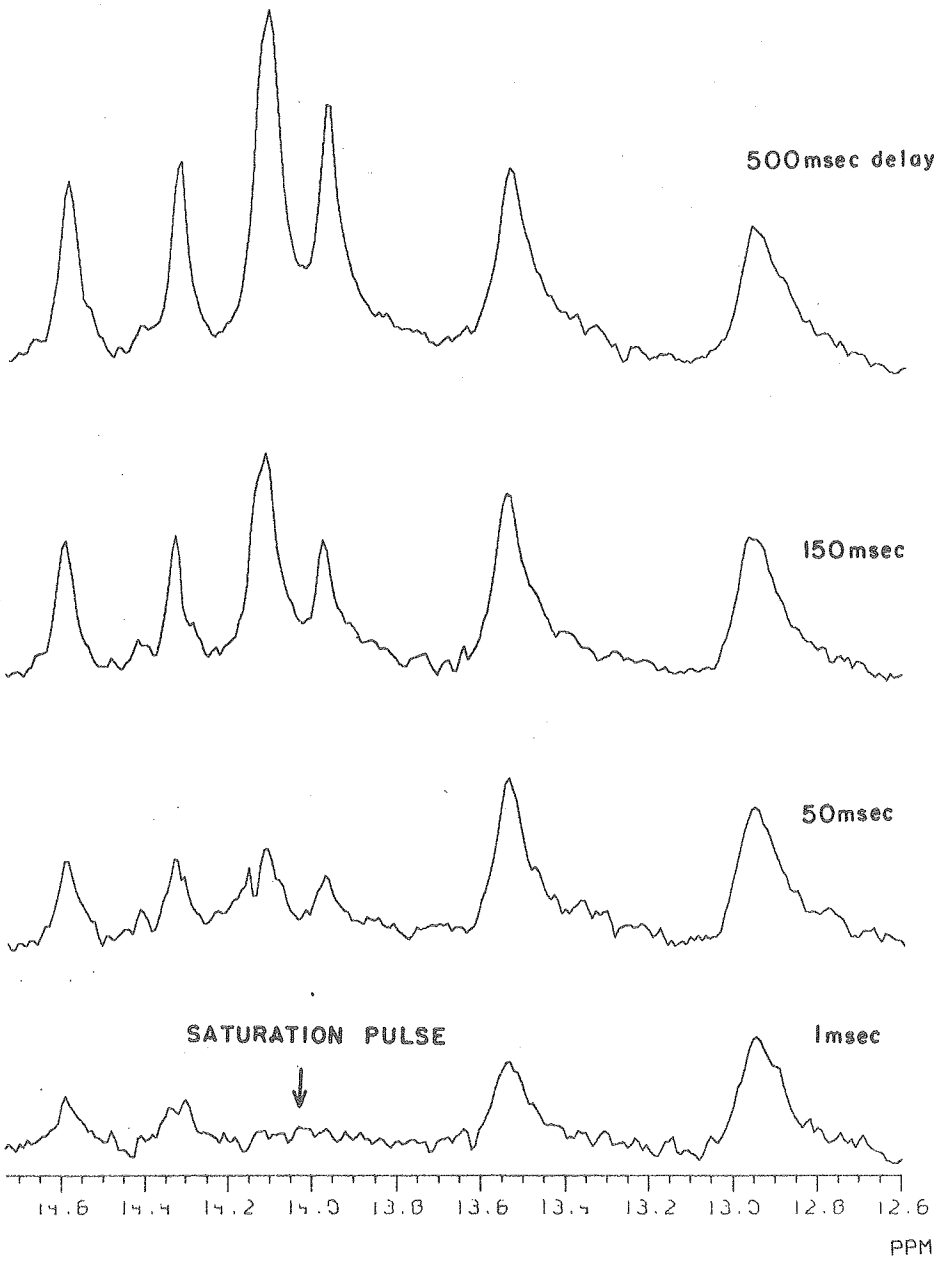
Saturation recovery experiments were performed on the imino protons of the oligonucleotide helices $dCA_5G + dCT_5G$ and $rCA_5G + dCT_5G$. The experiment involves the saturation of one or several of the imino protons followed by a variable delay time before taking the observation Redfield pulse. An example of the data for $dCA_5G + dCT_5G$ is shown in Figure 3.5. Here the saturation pulse is centered at 14.1 ppm and as the delay time increases, the intensity of the saturated peak also increases. The peak recovers due to exchange with unperturbed water and by other relaxation processes such as the spin-lattice relaxation of the imino protons. (This will be discussed in a later section.)

The rate of exchange is at least partially dependent upon the rate of opening of the helix under these conditions and since the rate of opening of the helix is highly temperature dependent, i.e. has a large activation energy, the exchange rates measured may also be temperature dependent. For this reason special precautions were taken to test the temperature of the sample while in the NMR probe. Temperature regulation on the NMR spectrometers is done by flowing a stream of N_2 gas past the sample and monitoring the temperature of the gas.[64] The thermocouple which monitors the gas is usually close to, but not at the same point as the sample tube. The temperature at the sample was obtained by measuring the chemical shift differences of neat methanol, which is

Figure 3.5. Saturation recovery experiment on $dCA_5G + dCT_5G$ double helix.

SATURATION RECOVERY EXPERIMENT USING
REDFIELD 214 PULSE SEQUENCE

270MHz SPECTRA OF $dCA_5G + dCT_5G$ AT 5°C IN H_2O



XBL 805-9523

calibrated with temperature. This procedure is not as accurate but is more convenient than the other method used which involves inserting a small thermocouple directly inside an NMR tube and then putting the tube in the probe. This gives a direct measure of the sample's temperature on a digital thermometer connected to the thermocouple. Both the long term (month to month) and the short term (time of experiment) stability and accuracy of both the 270 MHz and 360 MHz instruments were found to be very good, with an error of no more than ± 1°C, at fairly high N_2 flow rates.

3. RESULTS

A) Saturation Recovery Experiments

Relaxation rates for the imino protons were analyzed in the following way. The imino protons exchange between three different environments: the helix, the coil, and water as shown below.

$$H_h \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} H_c \qquad (1)$$

$$H_c + HOH^* \underset{k_{cw}[catalyst]}{\overset{k_{cw}[catalyst]}{\rightleftharpoons}} H_c^* + HOH$$

where H_h and H_c are the imino protons in the helix and coil respectively. Crothers et al.[56] derived the Bloch equations for exchange between three different environments, an extension

of McConnell's two state derivation.[65] The decay of M_z, the magnetization along the z axis, for each state is given below; with all the M_x's and M_y's equal to zero.

$$\frac{dM_{zh}}{dt} = -\frac{M_{zh}}{\tau_{op}} + \frac{M_{zc}}{\tau_{cl}} + \frac{M_{zh^o} - M_{zh}}{T_{1h}} \qquad (3)$$

$$\frac{dM_{zc}}{dt} = -\frac{M_{zc}}{\tau_{cl}} - \frac{M_{zc}}{\tau_{cw}} + \frac{M_{zh}}{\tau_{op}} + \frac{M_{zw}}{\tau_{wc}} + \frac{M_{zc^o} - M_{zc}}{T_{1c}} \qquad (4)$$

$$\frac{dM_{zw}}{dt} = -\frac{M_{zw}}{\tau_{wc}} + \frac{M_{zw}}{\tau_{cw}} + \frac{M_{zw^o} - M_{zw}}{T_{1w}} \qquad (5)$$

where M_{zh}, M_{zc}, and M_{zw} are the magnetization s along the z axis of the helix, coil, and water respectively; M_{zh^o}, M_{zc^o}, and M_{zw^o} are the equilibrium magnetizations of the helix, coil, and water respectively; and T_{1h}, T_{1c}, and T_{1w} are the longitudinal relaxation times in the helix, coil, and water, respectively. The lifetimes are: $\tau_{op} = 1/k_{op}$, $\tau_{cl} = 1/k_{cl}$[single strands] and $\tau_{wc} = \tau_{cw} = 1/k_{cw}$[catalyst]. This system is greatly simplified under our conditions by the fact that $1/\tau_{cw} = k_{cw}$[catalyst] $>> 1/T_{1c}$. This means there is essentially no relaxation due to the T_1 of the coil, because the coil is in such fast exchange with water: a proton spends so little time in the coil state that there is negligible relaxation from that state. A second, similar assumption is that $1/\tau_{cw} >> 1/\tau_{cl}$, which implies that everytime the helix opens to the coil state the protons exchange with water. These two assumptions allow a two-

state analysis of this system with just the helix and water states. One can now write the equilibrium as

$$H_h \underset{k_{wh}}{\overset{k_{hw}}{\rightleftharpoons}} \text{water} \tag{6}$$

and the modified Bloch equations as

$$\frac{dM_{zh}}{dt} = -\frac{M_{zh}}{\tau_{hw}} + \frac{M_{zw}}{\tau_{wh}} + \frac{M_{zh^\circ} - M_{zh}}{T_{1h}} \tag{7}$$

$$\frac{dM_{zw}}{dt} = -\frac{M_{zw}}{\tau_{wc}} + \frac{M_{zh}}{\tau_{cw}} + \frac{M_{zw^\circ} - M_{zw}}{T_{1w}} \tag{8}$$

In the saturation recovery experiment performed here water is not perturbed, so both equation (8) drops and $M_{zw} = M_{zw^\circ}$ in equation (7). The steady state condition of $M_{zh^\circ}/\tau_{hw} = M_{zw^\circ}/\tau_{wh}$ gives the following equation

$$\frac{dM_{zh}}{dt} = \frac{-M_{zh} + M_{zh^\circ}}{\tau_{hw}} + \frac{M_{zh^\circ} - M_{zh}}{T_{1h}} \tag{9}$$

which can be integrated with the initial condition of $M_{zh} = 0$ at $t = 0$ to

$$M_{zh}(t) = M_{zh^\circ}(1 - e^{-t[(1/\tau_{hw}) + (1/T_{1h})]}) \tag{10}$$

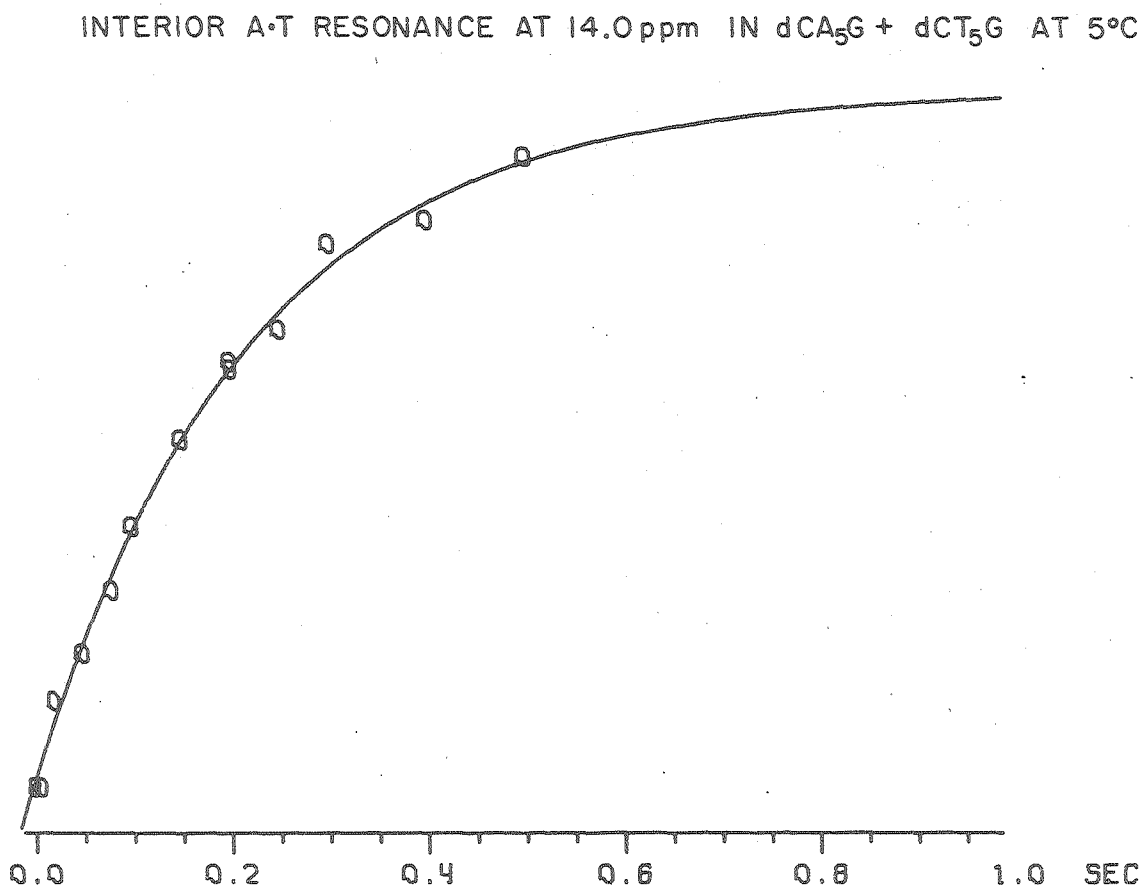
Data reductions were done on a Nicolet 1180 computer which performs a least squares fit on the above equation by changing M_{zh° and $[(1/\tau_{hw}) + (1/T_{1h})]$. The program then gives the best fit to the lifetime $[(1/\tau_{hw}) + (1/T_{1h})]$, and the equilibrium intensity of the peak $M_{zh}{}^\circ$. The experimental

intensities of the peaks were obtained by measuring their heights or areas, with areas used for broad resonances since this procedure seemed to produce more accurate results than the heights. The areas were measured by weighing traces of the peaks on an analytical balance.

An example of the fits are given in Figure 3.6. Experiments were usually done with 10-20 data points and never less than 8 points. At least one point was taken at a very short delay time, 0.25 or 0.5 msec, in order to estimate the extent of saturation at zero time. The power of the saturation pulse was usually set so as to get complete saturation at the shortest time, but peaks nearby may be only partially saturated. It is possible to get relaxation information from partially saturated peaks by subtracting the intensity of the peak at time zero from all the other delay times. A point was also usually taken where the peak was fully recovered in order to get an idea of the equilibrium intensity. This was done by using a delay time which is much larger than (4 - 5 times) the relaxation time, or taking a spectrum where no power was put in the saturation pulse. For a given experiment all the spectra had the same number of scans and were processed in the same way, so that their intensities could be directly compared. The error on the relaxation time of an individual calculation was on the order of 5 - 30% and dependent on the number of delay times measured and also the signal to noise of the individual spectra.

Figure 3.6. Fit of lifetime for saturation recovery experiment of $dCA_5G + dCT_5G$ helix at 5°C. The line shows the calculated curve and the circles show the experimental points.

FIT OF LIFETIME FOR SATURATION RECOVERY EXPERIMENT



XBL 8011-7494

B) Lifetimes of the Imino Protons for $dCA_5G + dCT_5G$ at 5°C

Saturation recovery experiments on the imino protons of $dCA_5G + dCT_5G$ at 5°C show varied lifetimes for protons in different parts of the helix. Table IX and Figure 3.7 show the lifetimes measured using equation (10) for this helix. The numbers given are the average of at least four measurements, with the errors representing the standard deviation of these measurements. No field dependence between the 270 MHz and 360 MHz instruments was seen for these lifetimes. Neither the length of the saturation pulse, nor the power of that pulse affected the lifetimes of the imino protons. The two peaks at 13.8 ppm (see Figure 3.7) are due to two interior A·T imino protons and were analyzed as one peak which recovered with one rate.

C) Temperature Dependence of the Relaxation Rates for $dCA_5G + dCT_5G$

The temperature dependence of the lifetimes for two of the protons in the deoxy helix were observed. One proton at 14.3 ppm was on an A·T base pair on the interior of the helix, while the other at 12.93 ppm was the terminal #1 C·G base pair. The temperature dependences are given in Table X. There are only three temperature points for the #1 C·G proton, while the interior A·T proton had four temperature points. The discussion of these data will be left to a later section.

TABLE IX

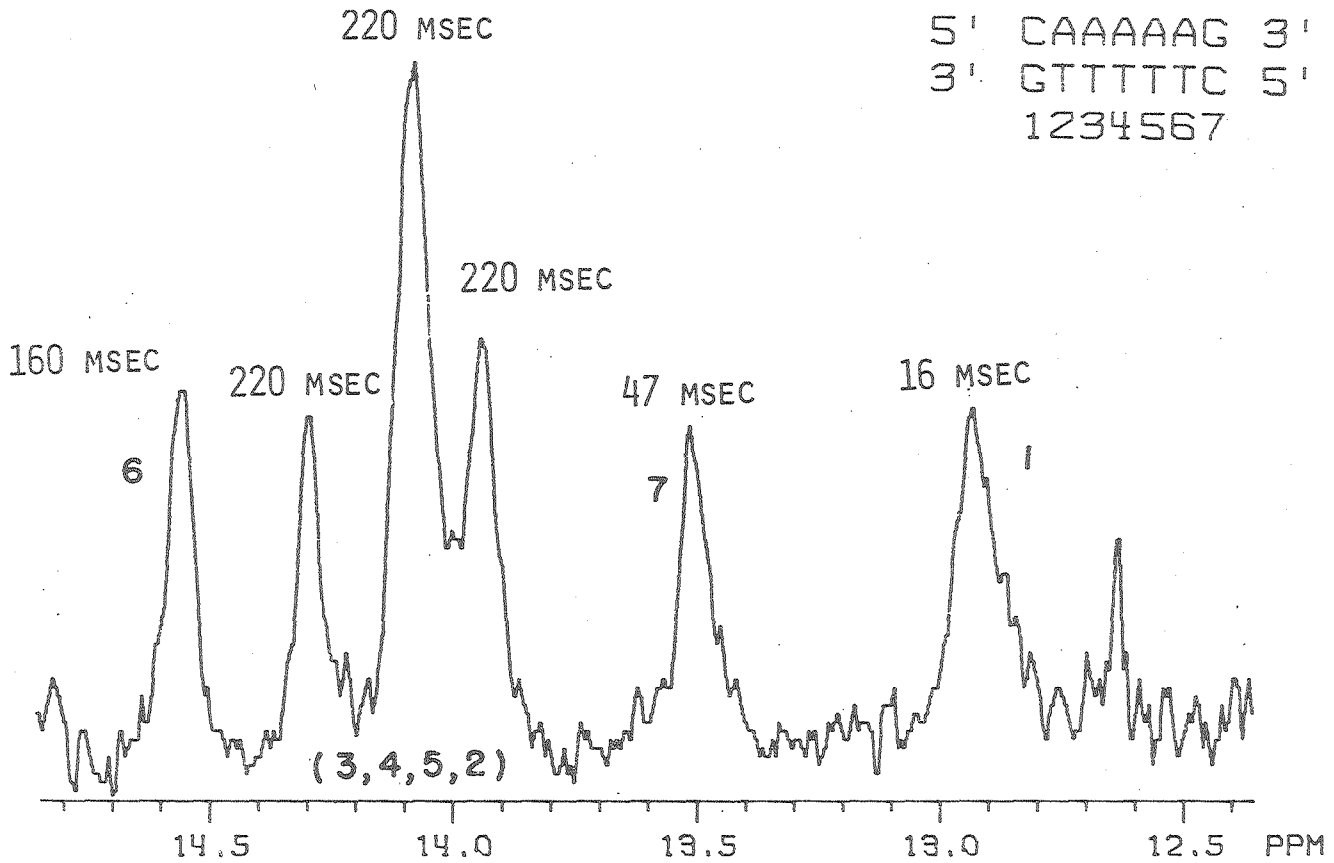
LIFETIMES OF IMINO PROTONS IN $dCT_5G + dCA_5G$

DOUBLE HELIX AT 5°C (msec)

chemical shift of peak	12.93 C·G(#1)	13.5 C·G(#7)	13.9 A·T(Int)	14.1 A·T(Int)	14.3 A·T(Int)	14.6 A·T(#6)
360 MHz (# of points)	16±2.4 (5)	47±10 (4)	208±20 (3)	237±20 (3)	220±18 (6)	173±20 (3)
270 MHz (# of points)	17 (1)	-	228 (2)	180 (1)	-	143±20 (2)
Average lifetimes using points at 360 and 270 MHz	16±2.2	47±10	217±20	220±31	220±18	161±26

Figure 3.7. Lifetimes of the imino protons in $dCA_5G + dCT_5G$ helix at 5°C.

LIFETIMES OF IMINO H-BONDING PROTONS IN THE HELIX $dCA_5G + dCT_5G$ AT 5°C



XBL 8011-7492

TABLE X

TEMPERATURE DEPENDENCE OF LIFETIMES IN $dCA_5G + dCT_5G$ (msec)

Temperature	Peak 12.93 C·G(#1)	Peak 14.3 (Interior A·T)
3°C	–	∿300
(# of points)		(1)
5°C	16	220
(# of points)	(5)	(6)
7°C	13	260
(# of points)	(1)	(1)
10°C	7	205
(# of points)	(2)	(2)
16°C	-	∿88
(# of points)		(1)

D) Lifetimes of the Imino Protons for the Hybrid Helix at 5°C

The lifetimes of the imino protons of the $dCT_5G + rCA_5G$ helix at 5°C are given in Table XI. There are fewer points than for the deoxy helix, so the errors on the lifetimes are on the order of 25%, except for the peaks at 13.4 and 13.5 ppm where the errors are slightly larger. These two peaks overlap slightly (see Figure 2.7) and, even though they have much different relaxation rates, it is still difficult to separate the contributions of the individual rates. The two peaks at 13.8 and 13.9 ppm relax with the same rate and were analyzed together.

4. DISCUSSION

The relaxation rates of the imino protons after saturation has been studied in several oligonucleotides.[62,63] There may be many processes involved in the relaxation of these protons. For example longitudinal relaxation will contribute to the rate as well as chemical exchange rates between unperturbed water molecules. Saturation recovery experiments cannot distinguish between these relaxation mechanisms. Other relaxation processes involving magnetic cross-relaxation mechanisms may also be important. For example, nuclear Overhauser effects between the imino protons and other oligomer protons or water protons can contribute to relaxation.[66] Johnston and Redfield have studied these effects on the relaxation of imino protons in tRNA, and saw two types of

TABLE XI

LIFETIMES OF IMINO PROTONS IN $rCA_5G + dCT_5G$ DOUBLE HELIX AT 5°C (msec)

Peak	12.25	13.5	13.4	13.8	14.0	14.6
360 Mhz (# of points)	∿22 (1)	9 (1)	63 (1)	-	-	-
270 MHz (# of points)	-	∿18	48	107	100	35
Average of lifetimes at 270 and 360 MHz	22	∿14	56	107	100	44

behavior.[62,63] At higher temperatures the relaxation seemed to be exchange dominated. This was deduced by the temperature dependence of the relaxation times and also the fact that no isotope effect was seen on the rates in the high temperature range. In going from 5 to 60% D_2O one would expect a large change in the efficiency of magnetic dipolar cross relaxation mechanism involving exchangeable resonances, because the efficiency of H-H dipolar relaxation is much larger than H-D relaxation. Johnston and Redfield saw a sizable isotope effect at low temperatures in the tRNA and concluded that in this range the relaxation was mainly due to dipolar relaxation mechanisms.[63] At higher temperatures, with faster rates, they saw little isotope effect, and also observed temperature dependences of the rates, from which they concluded that exchange was dominating the relaxation mechanism.

The interior A·T protons on this helix have lifetimes of 220 msec at 5°C (see Figure 3.7). There is no measurable temperature dependence of the lifetime from 3 - 10°C for the interior A·T at 14.3 ppm, as seen in Table X. It is not until 16°C that this lifetime changes appreciably, where it has a value of $\sim$ 90 msec. At this point exchange is probably a more important relaxation process than the dipolar relaxation. At temperatures below 10°C the measured lifetimes (220 msec) are not due to chemical exchange but other faster processes.

Thus the exchange lifetimes and lifetimes for opening of the helix are greater than 220 msec below 10°C. The terminal base pairs on the helix have much shorter lifetimes and show some temperature dependence, so these protons are probably in the exchange limiting range.

An interesting result seen in this oligonucleotide helix is the large difference in the lifetimes of the two terminal C·G base pairs. The C·G #1 imino proton has a lifetime of 16 msec, while C·G #7's lifetime is 47 msec, as seen in Figure 3.7. This effect is presumably due to the differences in the energy of breaking the H-bonding structure for the two ends of the helix. Different stacking energies for these two terminal base pairs are most likely the major contributions to the different opening rates.

For the hybrid helix $dCT_5G + rCA_5G$ the lifetimes of the imino protons are much shorter than the lifetimes in the deoxy helix at 5°C, as shown in comparison of Table IX and XI. The rates in this helix are more likely to be exchange rate dominated even for the interior A·T base pairs because of the lower stability of the hybrid helix.[44] This difference in rates is probably a reflection of the faster exchange rates in the hybrid helix. There are again differences in the lifetimes of the two terminal C·G imino protons, although not as large as in the deoxy helix. The C·G #1 imino proton has a lifetime of 22 msec while the C·G #7 has a lifetime of 14 msec. The interior A·T protons have lifetimes in the range of 100 msec.

A) Are These Helices in the Open Limited Region?

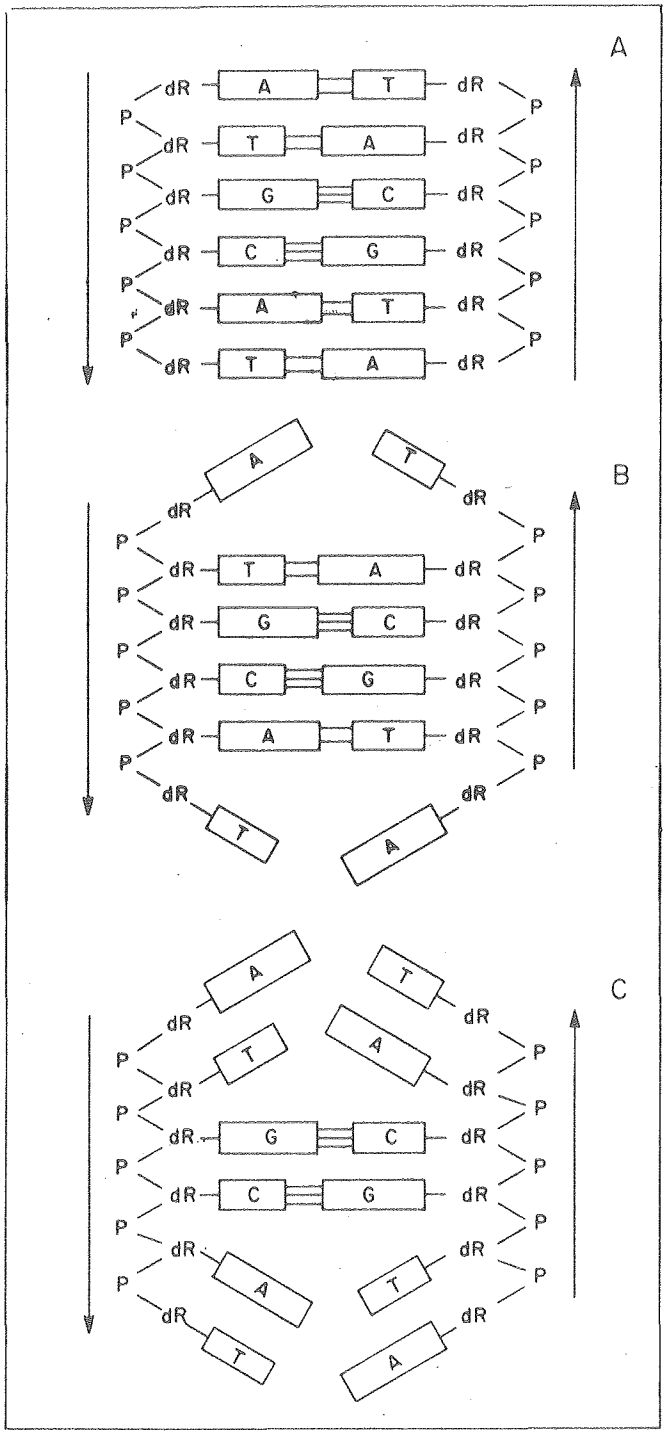
An assumption that exchange of the imino protons was open limited was made in the derivation used for analysis of the saturation recovery experiments. The validity of that assumption is based on other experimental work performed on the exchange of imino protons under similar conditions.[19, 53-55] The assumption states that $1/\tau_{cw} >> 1/\tau_{cl}$ (see equations 1 and 2) and implies that everytime the helix opens to the coil state, the protons exchange with water. Here $1/\tau_{cw} = k_{cw}[\text{catalyst}]$ where exchange is a base-catalyzed reaction, and the catalyst in our buffer is HPO_4^{2-}. The HPO_2^{2-} concentration was 11 mM under our conditions. The rate constant k_{cw} for HPO_4^{2-} has been measured in the oligonucleotide system dATGCAT to be 2×10^6 $M^{-1}sec^{-1}$ at 5°C, so $1/\tau_{cw} = 2.2 \times 10^4$ sec^{-1} in our system. We now want to compare this to $1/\tau_{cl}$ where $1/\tau_{cl} = k_{cl}[\text{single strand}]$, and k_{cl} for oligonucleotides has been measured by temperature jump methods on other oligonucleotides to be $\sim 10^6$ $M^{-1}sec^{-1}$ at this temperature.[19,20] $1/\tau_{cl}$ will then be less than 10^2 in our experiment, so in this case $1/\tau_{cw} >> 1/\tau_{cl}$, and our analysis is valid. This will be true for the interior protons which will exchange only when the whole helix goes from double to single strand.

This is the same result found by Hilbers and Patel for dATGCAT,[54-55] where the interior C·G base pairs exchange only when the whole double strand goes to single strand. The A·T imino protons were found to have much faster exchange rates than the double to single strand rate. Their system was modeled by the three states shown in Figure 3.8. The A·T imino protons can exchange by going through state B or C; but only by going to the single strand can the interior C·G imino proton exchange. The end base pairs thus open and close at much faster rates than the double to single strand rates, and this behavior is termed fraying on the ends of the helix. Kallenbach et al. observed very similar results in the helix dAAAGCTTT.[19]

In our helices dCA_5G + dCT_5G and rCA_5G + dCT_5G we also observe fraying on the ends of the helix, with this being manifested as shorter lifetimes of the terminal base pair imino protons relative to the interior protons in the helices. Our analysis assumes that even on the end base pairs, the exchange is open limited. This means that every time a base pair opens, the imino proton exchanges. The available evidence cannot show that this assumption is true, but we think that it is more likely than in the case of dATGCAT or AAAGCTTT because of the more stable C·G base pairs on the ends of our oligonucleotide helices. Evidence that opening was not the rate limiting step for exchange of the A·T imino protons in dATGCAT and dAAAGCTTT was mainly from the chemical

Figure 3.8. Proposed structures of dATGCAT in double helix (from reference 73).

Scheme I.



XBL 8011-7465

shift of the imino protons with temperature.[19,54,55] They found fairly large chemical shift changes for the A·T base pairs, up to 0.25 ppm for the terminal A·T imino proton in dATGCAT, but little chemical shift changes for the internal C·G imino protons. The chemical shift changes of the A·T base pairs are from the closed and open state being in fast exchange. Therefore the shift of the imino resonance with temperature is from changes in the populations of the two states with temperature, and not due to exchange with H_2O. In our helices we see very little chemical shift changes of the imino protons with temperature. Figures 2.9 and 2.17 show that there are almost no chemical shift changes of the terminal, or interior, base paired imino protons in the dCA_5G + dCT_5G helix or the rCA_5G + dCT_5G helix, until the temperature where the resonances are very broad. So in our systems the assumption of open limited exchange is much more likely than in oligonucleotides with A·T base pairs on the ends. Since there are no chemical shift changes of the imino protons until they are extremely broad in these helices, the analysis used by Patel and Hilbers in dATGCAT is not applicable.[54] Thus the only reasonable interpretation of the data that can be made is the assumption of open limited exchange for all the imino protons.

5) CONCLUSIONS

Studies on the exchange rates of the imino protons of $dCA_5G + dCT_5G$ and $rCA_5G + dCT_5G$ were done by proton magnetic resonance. For the deoxy helix the results showed that the terminal base pairs have much faster exchange rates, and open much more often than the internal base pairs. The two terminal C·G base pairs in this helix also have different exchange rates, indicative of a sequence dependence for these exchange rates. One also expects that the rate of opening of the terminal base pair in a helix is essentially independent of the length of the helix. So that a C·G base pair on the end of the helix $\frac{CAAAAAA}{GTTTTTT}$ would open at the same rate even if this sequence were added to the end of a much longer more stable polynucleotide helix. The stability of the rest of the helix is not the rate determining factor in exchange on the ends, but only the sequence of the neighboring base pairs.

The interior protons of the $dCA_5G + dCT_5G$ show lifetimes at 5°C which are most likely due to other relaxation processes besides exchange with water. What these lifetimes do provide is a lower limit on the lifetime of the helical state, with this lifetime being greater than 220 msec at 5°C.

For the hybrid helix $rCA_5G + dCT_5G$ the measured lifetimes of the terminal base pairs are again much shorter than those of the interior base paired imino protons, and also show a slight sequence dependence. The lifetimes of the interior A·T imino protons are a factor of 2 shorter than the interior

A·T protons in the deoxy helix at 5°C. This is indicative of a lower stability of the hybrid helix compared to the deoxy helix. Whether this shorter lifetime for opening of the hybrid helix compared to the deoxy helix is a property of this particular sequence of nucleic acid, or is general for all sequences, is a question still to be answered.

PART II

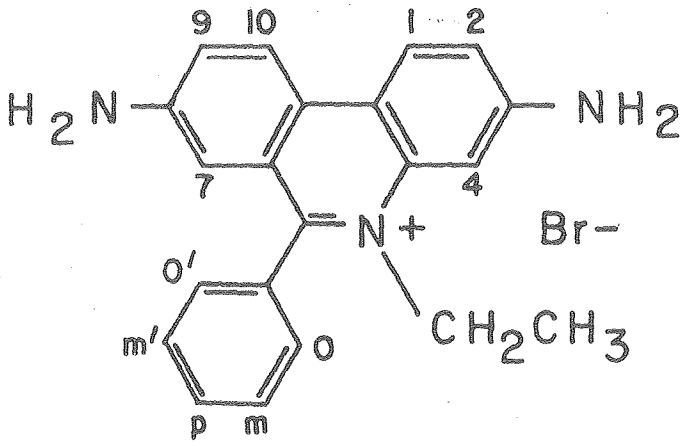
CHAPTER IV

1. INTRODUCTION

Ethidium (Et.) is a planar aromatic cation (shown in Figure 4.1) known to intercalate into DNA and RNA.[68] It has been shown to be a frameshift mutagen in the Ames test.[11] Spectroscopic studies of ethidium (and other intercalating drugs) binding to nucleic acids have recently been reviewed.[39,69] Krugh et al. have extensively studied the physical properties of ethidium binding to oligonucleotides by NMR, circular dichroism (CD), UV-visible absorption, and fluorescence methods.[72,77] Their studies on dinucleoside monophosphates with ethidium have shown the formation of mini-helices, with ethidium intercalated between the two base pairs. X-ray studies have provided detailed structures of ethidium in complexes with iodo rCpG and iodo rUpA.[70,71] A sequence dependence for ethidium binding has been found where ethidium preferentially binds to pyrimidine (3' - 5') purine sites compared to purine (3' - 5') pyrimidine sites.[72]

Proton magnetic resonance studies on dinucleoside monophosphates with ethidium[73] and a similar analog propidium[74] have shown that the same mixed sugar pucker seen in the crystal complexes also exists in solution. The sugar

Figure 4.1. Structure of ethidium bromide.



Ethidium Bromide

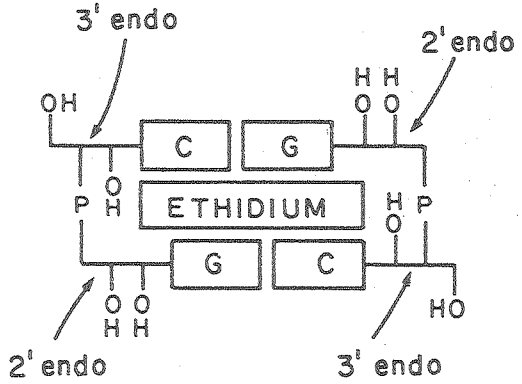
XBL 782-7422

puckers found in solution are 3' endo for the 5' residue and 2' endo for the 3' residue in the molecule. This is indicative of at least partial similarity between the crystal and solution structures.

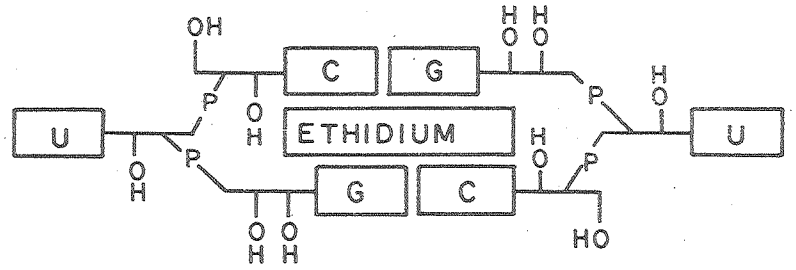
Lee and Tinoco have studied complexes of rCpG, rCpUpG, and rGpUpG + rCpC with ethidium by proton NMR.[73] These systems represent intercalation of ethidium into a perfect helix (rCpG), a helix with a bulged uracil on both strands (rCpUpG), and a helix with bulge on only one strand, (rGpUpG + rCpC). The proposed structures for these complexes are schematically shown in Figure 4.2. Bulged base structures may be important in the mechanism of frameshift mutagenesis by ethidium. Specifically ethidium may intercalate in transient bulges formed during replication, repair, or recombination of DNA.[73]

In this chapter we will investigate the thermodynamics of ethidium binding to several different dinucleoside monophosphates (rCpG, rUpA, and dCpG), a trinucleoside diphosphate which can form a bulge on both strands (rCpUpG), and a mixture of a trinucleoside and a dinucleoside (rGpUpG + rCpC) which can forma bulged structure on one strand. We will try to answer questions about the properties governing the sequence specificity of ethidium, and the strength of ethidium binding to perfect helices compared to those with bulges.

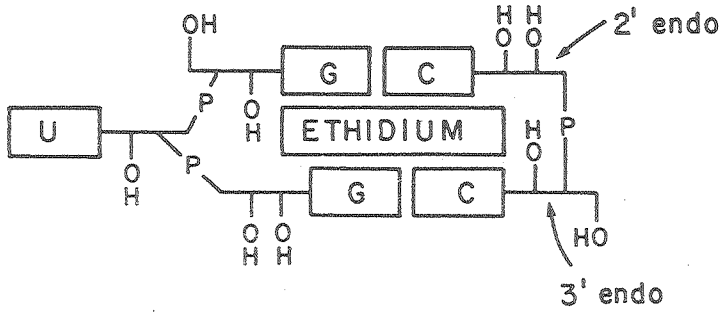
Figure 4.2. Proposed structures for ethidium binding to oligonucleotides (from reference 73).



CpG



CpUpG



GpUpG + CpC

XBL 782-7157

2. OPTICAL STUDIES EXPERIMENTAL

A) Materials

The ribo- dinucleoside 3'-5' monophosphates were purchased from Sigma. The deoxyribo- dinucleoside 3'-5' monophosphates were purchased from Collaborative Research. Ethidium bromide was obtained from Sigma; the trimers rGpUpG and rCpUpG were bought from Boehringer Mannheim, however the majority of the rCpUpG used in this study was enzymatically synthesized with polynucleotide phosphorylase using standard procedures.[24] The E. coli DNA used in this study was obtained from Worthington.

B) Methods

Visible absorption studies were done on a Cary 118 spectrometer, with an external bath (Haake model F) used for temperature regulation to an accuracy of ± 0.5°C. Melting studies were done on a Beckman DU spectrophotometer and a Gilford Model 2000 recorder, which measures temperature to ± 0.5°C. Circular dichroism spectra were measured on a Model 60 Cary spectropolarimeter equipped with a 6001 CD attachment. Equilibrium sedimentation studies were done on a Model E analytical ultracentrifuge, with an A_nF rotor.

All samples contained 8.0 mM Na_2HPO_4, 20 mM NaH_2PO_4, 0.18 M NaCl, 0.1 mM Na_2EDTA, pH = 7.0 unless otherwise noted. Centrifugation of the samples helped eliminate dust

particles, and was performed before most of the optical experiments. Extinction coefficients used for the dimers were from the CRC handbook.[29] The extinction cofficients of the trimers were calculated with the assumption of only nearest neighbor interaction, using the extinction coeffecients of the dinucleoside monophosphates and mononucleotides. The extinction coefficient of ethidium was taken as 5.45×10^3 at 480 nm.[86] The extinction coefficient of E. coli DNA was taken as 6.4×10^3 at 260 nm.[87] All these extinction coefficients are for 25°C.

The low salt experiments performed on E. coli DNA were done after exhaustive dialysis of the DNA against a solution of double distilled water with 0.6 mM NaOH added to adjust the pH of the water to 7.0. Thus the dialysate had a sodium ion concentration of 0.6 mM, but the DNA solution's sodium ion concntration was not known since the number of sodium counter ions bound to the DNA was not taken into account. The concentration of DNA, per phosphate, was around 0.5 mM so the actual sodium ion concentration in the DNA solution was probably less than 1.5 mM.

The ethidium concentration in most of the optical experiments was around 0.05 mM. The oligonucleotide concentrations ranged from 0.1 - 5 mM depending upon the experiment. Samples were prepared by added given amounts of ethidium and oligonucleotides from stock solutions, so the total concentrations of ethidium and nucleotide in each sample were known.

3. RESULTS

A) Determination of Equilibrium Constants of Ethidium Binding to Oligonucleotides

Equilibrium constants were calculated using the large change in the visible absorption spectrum of ethidium upon binding to nucleic acids.[75] The shift in the visible absorption is seen in Figure 4.3, which shows ethidium binding to the dimer rCpG at 1°C. The largest change in the

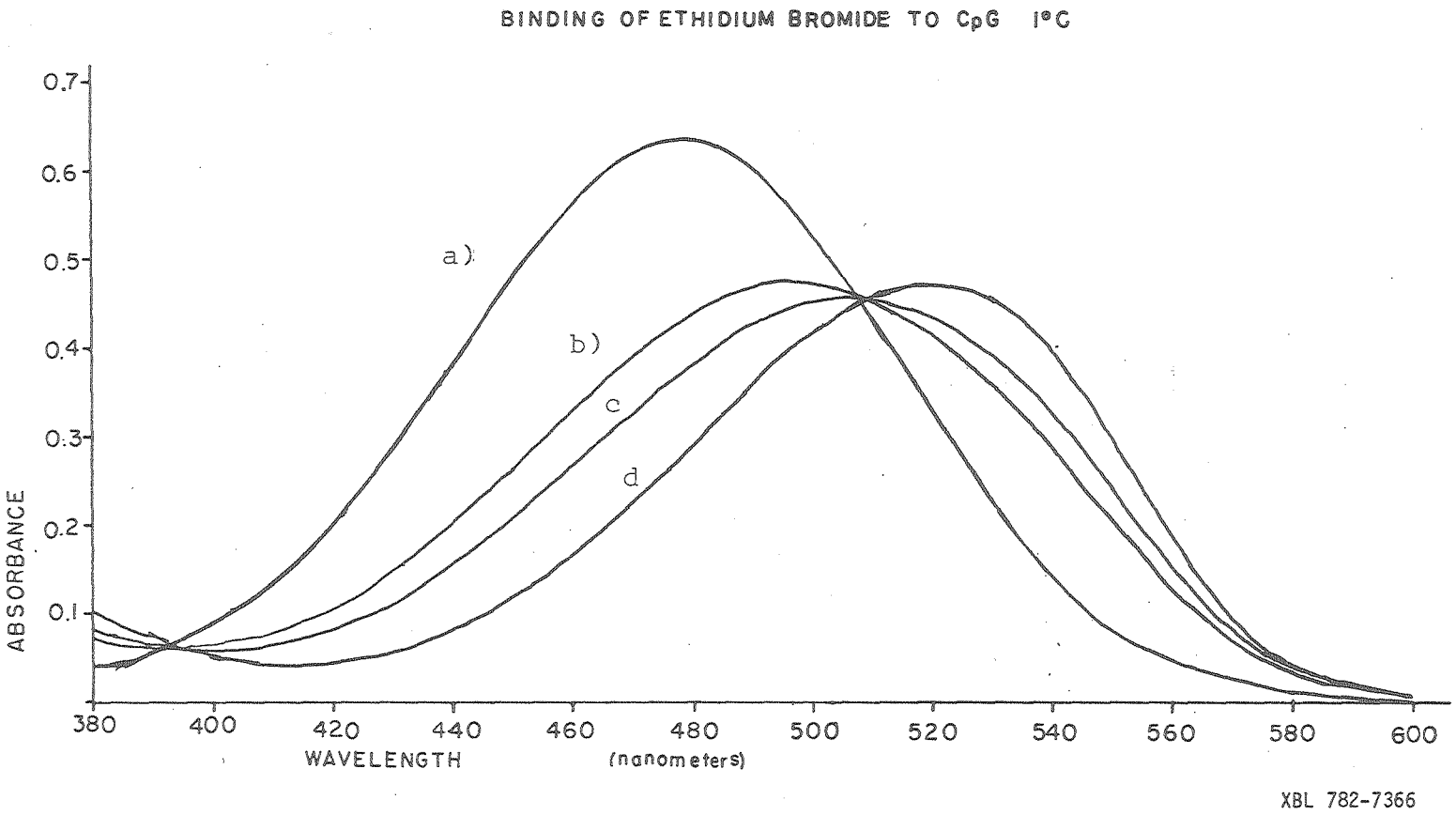
Figure 4.3. Binding of ethidium at different concentrations of rCpG. Ethidium concentration was constant at 0.12 mM. All solutions were run in standard buffer. The curves are:

a) free ethidium.

b) [rCpG] = 0.22 mM.

c) [rCpG] = 0.33 mM.

d) [rCpG] = 3.0 mM.


BINDING OF ETHIDIUM BROMIDE TO CpG 1°C
ABSORBANCE
0.7
0.6
0.5
0.4
0.3
0.2
0.1
a)
b)
c
d
380
400
420
440
460
480
500
520
540
560
580
600
WAVELENGTH
(nanometers)
XBL 782-7366

absorbance is at 465 nm and was the wavelength used in these calculations. The equilibrium constant for a 2 to 1 complex can be written as follows (the determination of the stoichiometry will be discussed in a later section)

$$Nuc_a + Nuc_b + Et. \rightleftharpoons complex \qquad (1)$$

$$(2) \qquad K = \frac{[C_{complex}]}{[C_{Nuc_{a^\circ}} - C_{complex}][C_{Nuc_{b^\circ}} - C_{complex}][C_{Et^\circ} - C_{complex}]}$$

where $C_{complex}$ is the concentration of the complex in solution and $C_{Nuc_a^\circ}$, $C_{Nuc_b^\circ}$ and C_{Et° are the total concentrations of nucleotide a, nucleotide b, and ethidium respectively, in solution. The analysis of the absorption data was done by two methods. One method used by Blake and Peacocke involves obtaining the spectrum of the dye in the complex.[76] This is done by holding the concentration of the dye constant and titrating with nucleic acid until all the dye is bound; with this spectrum and the spectrum of the free dye it is possible to obtain the concentrations of the free and bound dye. This analysis assumes that there are only two states for the dye in the system, either free in solution or bound to nucleic acid. It also assumes that the nucleic acid does not absorb at this wavelength. The method requires fairly high equilibrium constants in order to obtain a spectrum of the fully bound. If the total concentrations of the nucleic

acid and dye in solution are known then the equilibrium constant for dye binding can be calculated. Let

$$C_{complex} = \alpha C_{Et^\circ} \qquad (3)$$

$$\alpha = \frac{A_f - A}{A_f - A_b}$$

where A_f, and A_b are the absorbances of the free and bound dye respectively, A is the absorbance of an intermediate state where α, the fraction of ethidium in the complex, is being calculated. Table XII shows the calculations of the equilibrium constant for rCpG + ethidium at 0°C using this method.

The other method of analysis involves using a Benesi-Hildebrand type analysis. We can write

$$C_{complex} = \frac{A - \varepsilon_f C_{Et.^\circ} \ell}{(\varepsilon_b - \varepsilon_f)\ell} \qquad (4)$$

where A is the measured absorbance of the solution (normalized to 1 cm), $C_{Et.^\circ}$ the total concentration of ethidium, ℓ the pathlength of the cell in cm, ε_f the extinction coefficient of free ethidium at 465 nm, and ε_b the extinction coefficient of the ethidium in the complex. Substituting equation (4) into equation (2) and rearranging with the assumptions that $Nuc_{a^\circ} = Nuc_{b^\circ}$ and $Nuc^\circ = Nuc_{a^\circ} + Nuc_{b^\circ}$ we obtain (which is the form of the equation for self-complementary oligonucleotides)

TABLE XII

EQUILIBRIUM CONSTANTS FOR ETHIDIUM BINDING TO rCpG 0°C

CALCULATED USING PEACOCKE-TYPE ANALYSIS

A_b = 0.074 A_f 0.219 λ = 465 nm $\alpha = (A_f - A)/(A_f - A_b)$

[Nuc] (mM)	α	K (M^{-2})
1.3×10^{-4}	0.47	1.0×10^{8}
2.0×10^{-4}	0.67	9.3×10^{7}
2.7×10^{-4}	0.80	8.5×10^{7}
4.0×10^{-4}	0.91	8.6×10^{7}

average K = 9.1×10^{7} (M^{-2})

$$\frac{C_{Et^o}}{(A - \varepsilon_f C_{et^o}\ell)} = \frac{1}{(C_{Nuc^o} - 2C_{complex})^2(\varepsilon_b - \varepsilon_f)K}$$

$$+ \frac{1}{(\varepsilon_b - \varepsilon_f)} \qquad (5)$$

A plot of $C_{Et^o}/(A - \varepsilon_f C_{Et^o}\ell)$ versus $1/(C_{Nuc^o} - 2C_{complex})^2$ yields K and $(\varepsilon_b - \varepsilon_f)$ immediately. However, the quantity $(C_{Nuc^o} - 2C_{complex})$ is not known, but only the total concentration of the nucleotide C_{Nuc^o}. Assuming that the concentration of complex was small compared to the total concentration of added nucleotide, the first fitting to equation (5) had $C_{complex}$ set equal to zero. The resulting equilibrium constant was then used to find the value for $C_{complex}$ by substitution into equation (2). Refitting of the data in equation (5) with this value of $C_{complex}$ included gave a new estimate of the equilibrium constant. This iteration was done until convergence of the equilibrium constant occurred to within 1%. Figures 4.4a and 4.4b show ethidium binding to several dimers using this analysis. The equilibrium constants derived from these data will be given in a later section. Both the Peacocke and Benesi-Hildebrand methods gave equivalent results in our systems.

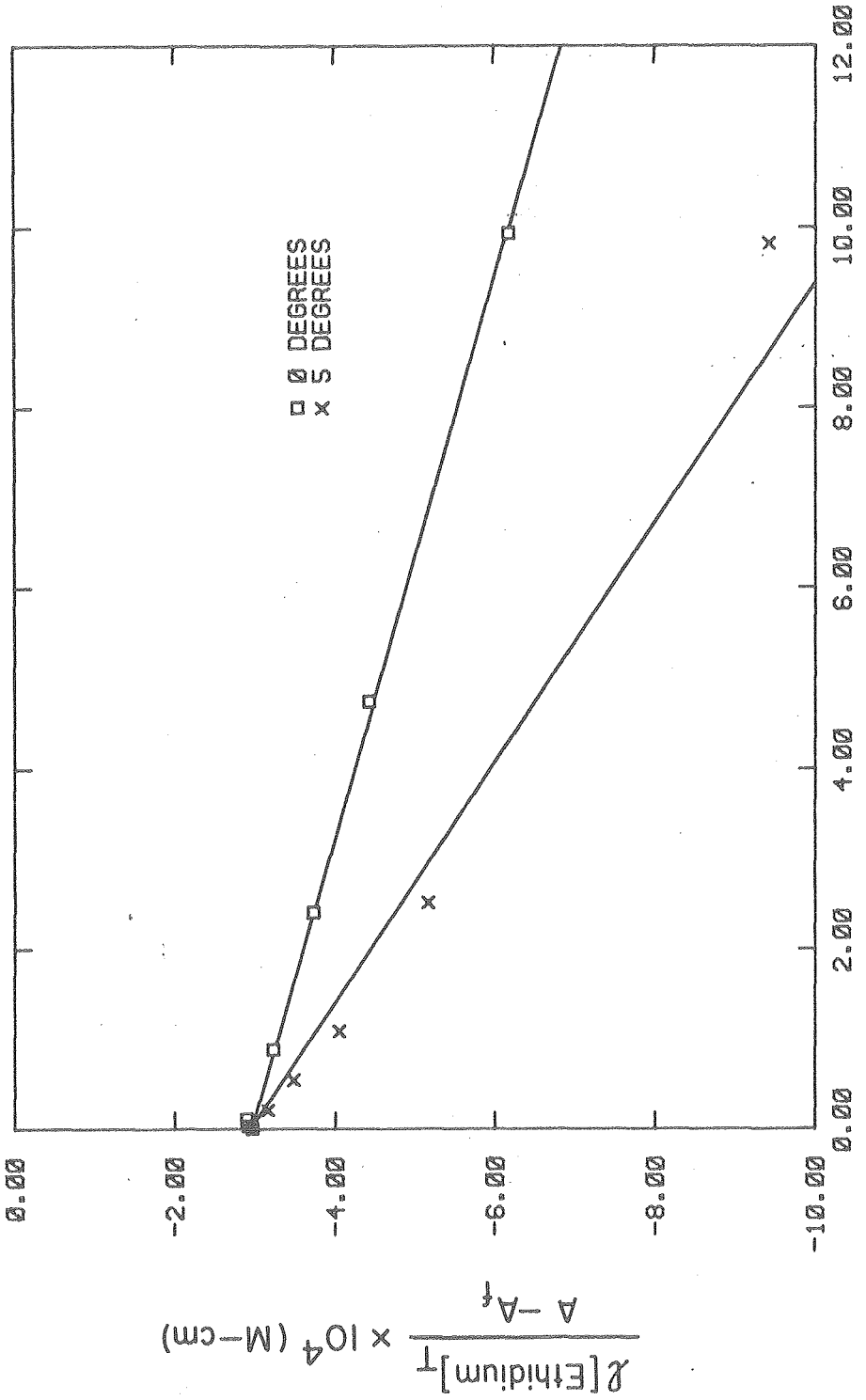
Figure 4.4.

a) Benesi-Hildebrand type plot for rCpG at 0 and 5°C.

b) Benesi-Hildebrand type plot for dCpG at 0 and 5°C.

c) Melting of ethidium with rCpG, for several concentration of ethidium and rCpG. The T_m for each concentration is at relative absorbance equal to 0.5.

a)

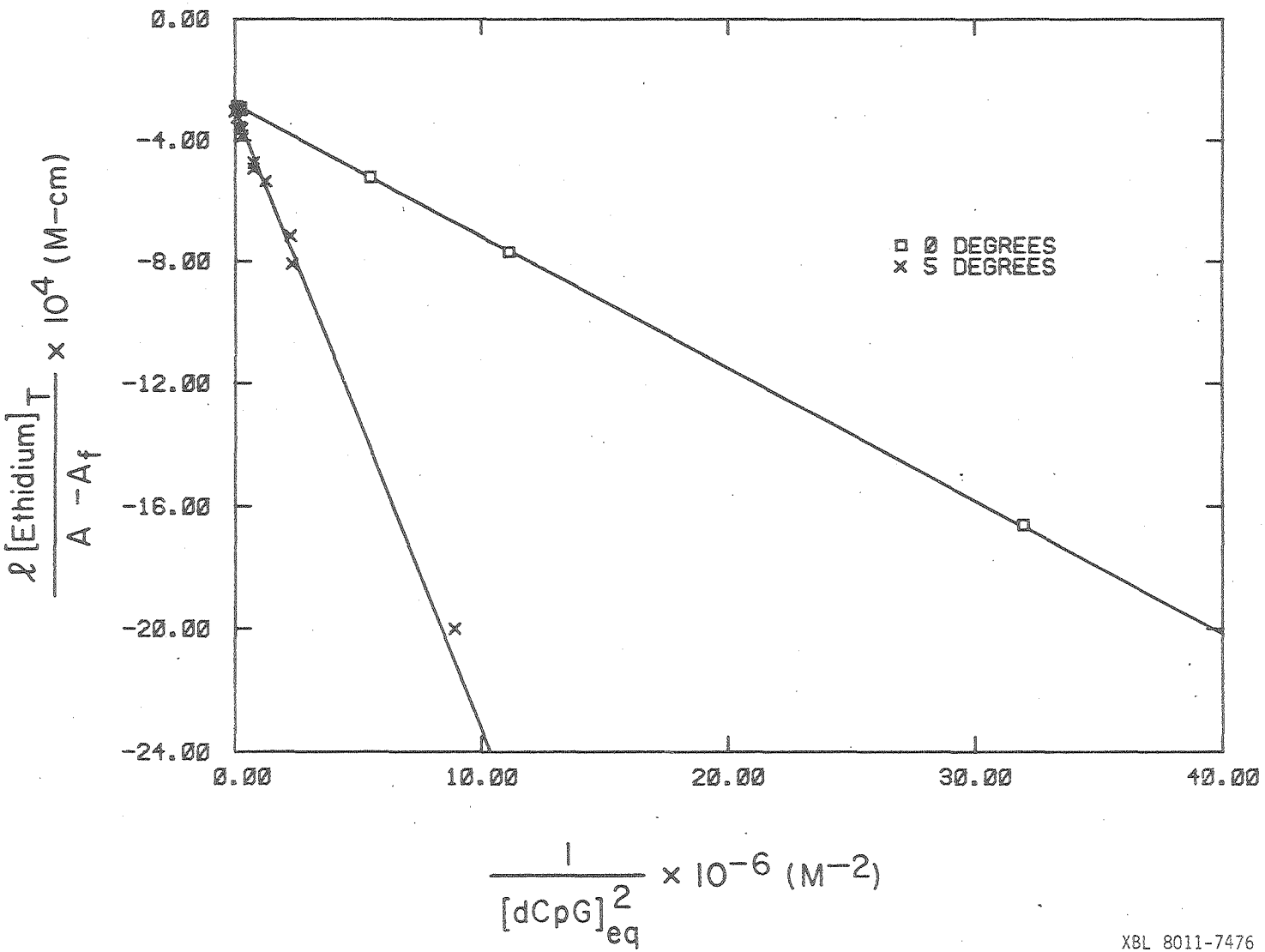


$$\frac{\beta[\text{Ethidium}]_T}{A - A_f} \times 10^4 \ (M^{-1}cm)$$

$$\frac{1}{[CpG]^2_{eq}} \times 10^{-7} \ (M^{-2})$$

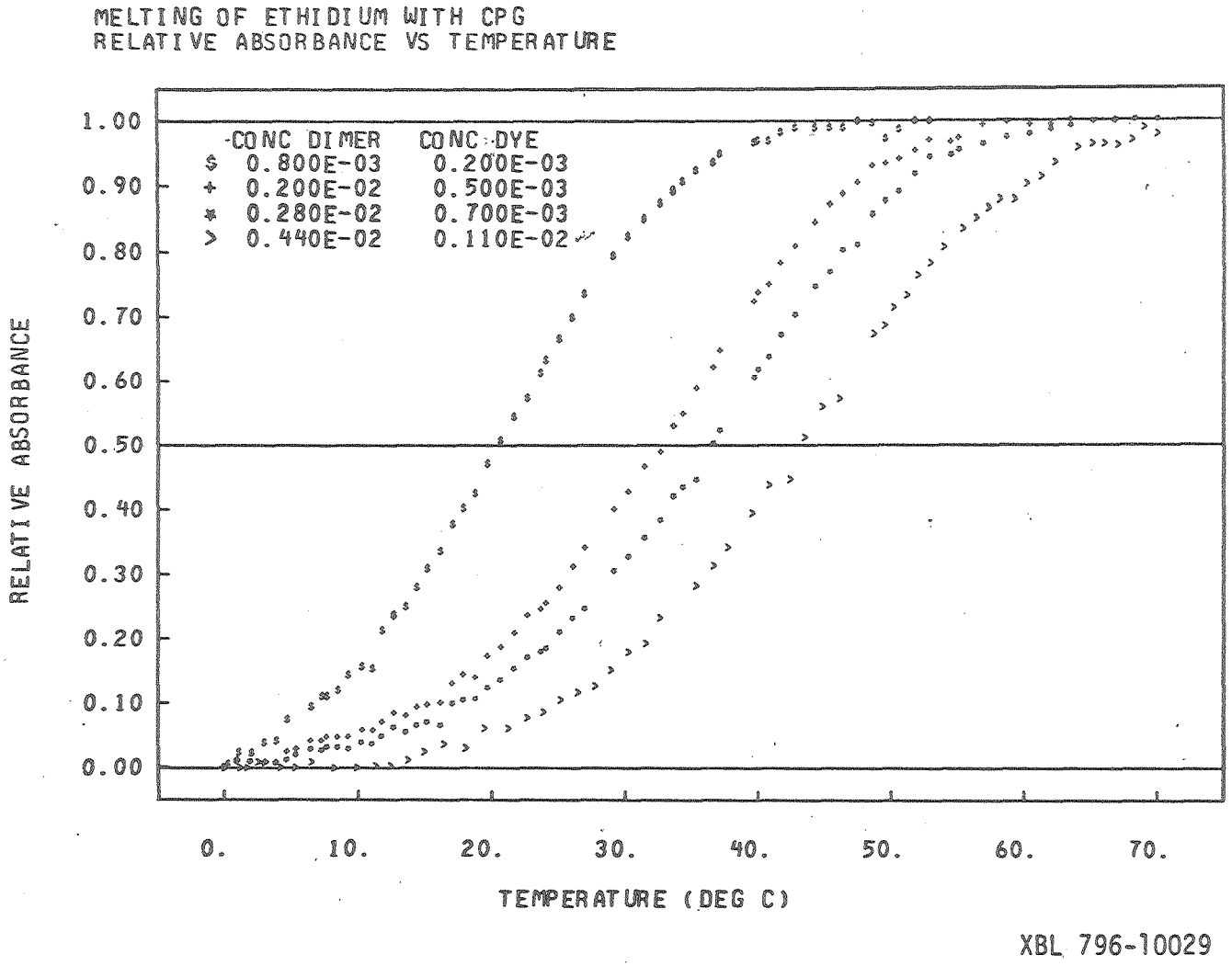
XBL 8011-7477

b)



XBL 8011-7476

c)



XBL 796-10029

This procedure worked well in our systems because the concentration of nucleotides was always well in excess of the ethidium concentration (usually 4 to 100 times the ethidium concentration). This also helped insure a 2:1 complex as opposed to the 2:2 complex seen in crystal structures of dimers with ethidium.[70,71]

B) Stoichiometry of the Complex

Krugh found complexes of dinucleotides and ethidium had stoichiometries of 2:1 dimer:ethidium under conditions similar to those in this work.[72,77] Our evidence for a 2:1 complex stems from the inability to fit the absorbance data of ethidium to other stoichiometries. There are good isosbestic points in the visible spectra of ethidium titrated with dinucleotides, which indicates there are only two states for the ethidium, free in solution, and bound in a complex. A Benesi-Hildebrand equation similar to (5) can be derived for a 1:1 complex of dimer with ethidium. The data do not fit a 1:1 complex, shown by the nonlinearity of the data fit using the 1:1 equation. A 2:2 complex was also ruled out in the concentration range used in this study. Inability of the data to fit a 2:2 stoichiometry was shown by assuming a Peacocke-type analysis for that equilibrium. The spectrum of the fully bound ethidium can be found by going

to very high nucleotide concentrations. The amount of ethidium in the complex for an intermendiate state (where the ethidium is parially bound and partially free) is found from equation (3). Knowing the concentration of the complex and the total concentrations of ethidium and oligonucleotide, it is possible to calculate the equilibrium constant. Each intermediate point in the titration curve will give a value for the equilibrium constant. These equilibrium constants should be the same over the whole titration curve. A large variation in the calculated equilibrium constant implies that the assumed stiochiometry is incorrect. Table XIII shows the equilibrium constants calculated for a 2:2 and a 2:1 complex for the titration of ethidium with rCpG at 0°C. The equilibrium constants for the 2:1 complex differ by about 10% over the titration, whereas this difference is over three orders of magnitude for the 2:2 complex. Clearly, the stiochiometry is 2:1 and not 2:2 for the complexes in the concentration range we are working.

C) Equilibrium Constants for Ethidium Binding to Oligonucleotides

Equilibrium constants for ethidium binding in a 2:1 complex were found for the following self-complementary oligonucleotides: rCpG, dCpG, rUpA, and rCpUpG. An upper limit on the equilibrium constant of rGpUpG + rCpC was also found. The equilibrium constants at 0°C are given

TABLE XIII

FIT OF DATA TO 2:1 AND 2:2 NUCLEOTIDE:ETHIDIUM COMPLEXES

rCpG at 0°C

$A_b = 0.074$ $A_f = 0.219$ $\lambda = 465$ $\alpha = (A_f - A)/(A_f - A_b)$

[Nuc] (mM)	α	calculated K for 2:1 complex (M^{-2})	calculated K for 2:2 complex (M^{-3})
1.3×10^{-4}	0.47	1.0×10^{8}	5×10^{14}
2.0×10^{-4}	0.67	9.3×10^{7}	7×10^{12}
2.7×10^{-4}	0.80	8.5×10^{7}	1×10^{12}
4.0×10^{-4}	0.91	8.6×10^{7}	3×10^{11}

below:

Nucleotide	Equilibrium Constant at 0°C (M^{-2})
rCpG	$100\pm20 \times 10^6$
dCpG	$8\pm2 \times 10^6$
rUpA	$1\pm0.3 \times 10^6$
rCpUpG	$1\pm0.3 \times 10^6$
rGpUpG + rCpC	$<10^5$

The errors in the equilibrium constants were obtained from errors in calculating the slope of the Benesi-Hildebrand type plots (see equation 5; Figures 4.4a and 4.4b show examples of these plots). Equilibrium constants at 5°C were obtained on several of these systems and are shown below:

Nucleotide	Equilibrium Constant at 5°C (M^{-2})
rCpG	$42\pm15 \times 10^6$
dCpG	$1.5\pm0.2 \times 10^6$

D) Binding of Ethidium to Single Strands

To find the strength of binding of ethidium to a single strand nucleotide in a 1:1 complex, the equilibrium constants for ethidium and rUpG or rCpA were measured. rUpG (or rCpA) is not selfcomplementary so it cannot form base pairs with itself in solution; this also rules out 2:1 dimer:ethidium complexes in solution. The equilibrium constant for the single strands were measured in the same manner as with the selfcomplementary nucleotides. An

equation similar to equation (5) was derived for 1 to 1 binding. The equilibrium constants obtained at 0°C are:

Nucleotide	K at 0°C (M^{-1})
rUpG	130 ± 30
rCpA	300 ± 100

These numbers are small compared to the binding of ethidium in the 2:1 complexes; thus single strand binding can be neglected in these systems.

E) Determination of the Enthalpy for Ethidium Binding to Oligonucleotides

$\Delta H°$ values for ethidium binding to these oligonucleotides were obtained from melting curves using the analysis of Martin et al.[24] By monitoring the absorbance of the ethidium at 465 nm it is possible to follow the amount of ethidium bound with temperature throughout the melt. Typical melting curves are shown in Figure 4.4c, for several concentrations of rCpG with ethidium. The ratio of total nucleotide to ethidium was held constant at 4:1 in all these experiments. The melting temperature of the ethidium, the T_m, is the point where half the ethidium is bound in the complex and half is free in solution. Equation (2) can then be used to calculate the equilibrium constant at this temperature. The slope of a plot of ln(K) versus $1/T_m$ is then $-\Delta H°/R$, which gives the enthalpy for complex formation.

Another method for calculating ΔH° is from the slope of the melting curve ($d\alpha/dT$) at the T_m, the formula derived for the case of 4:1 nucleotide:ethidium is[24]

$$\Delta H^\circ = -\frac{16}{3} R(T_m)^2 \left(\frac{d\alpha}{dT}\right)_{T_m} \qquad (6)$$

where α is the fraction of ethidium in the complex.

The ΔH°'s were calculated by both methods for these systems, and are given below along with the calculated ΔS° values:

Nucleotide	ΔH° ($d\alpha/dT$) (kcal/mole)	ΔH° ($\ln(K)$ vs $1/T_m$) (kcal/mole)	ΔS° (e.u.)
rCpG	-33	-32	-84
dCpG	-27	-29	-69
rCpUpG	-29		-79

F) Equilibrium Sedimentation Studies

Equilibrium sedimentation was done to determine whether aggregates exist in solution. What we were testing was to see if there was any aggregation of the 2:1 complexes with each other. This will be an important consideration in the studies of the induced CD of ethidium with dimers described later. Samples were run at 5°C and reached equilibrium after 24 hours. The weight average molecular weight of the complex was calculated using the following equation:[78]

$$M_{complex} = \frac{1}{2} M_{NaCl} Z\left(\frac{1 - \bar{V}_{NaCl}\rho}{1 - \bar{V}_{complex}\rho}\right) + Z\left[\frac{2RT \frac{\partial \ln(C)}{\partial r^2}}{(1 - \bar{V}_{complex})\omega^2}\right] \quad (7)$$

where Z is the charge on the complex, $\bar{V}_{complex}$ and $\bar{V}_{NaCl}$ are the partial specific volumes of the complex and sodium chloride respectively, ρ is the density of the buffer solution, C is the concentration of the complex at distance r from the axis of rotation, and ω is the angular velocity of the rotor. $\bar{V}$ for the complex was calculated by assuming it is the weight average of the component partial specific volumes. Thus for a complex of 2 CpG:1 ethidium, $\bar{V}$ is

$$\bar{V}_{complex} = \frac{2M_{CpG}\bar{V}_{CpG} + M_{Et}\bar{V}_{Et}}{2M_{CpG} + M_{Et}} \quad (8)$$

The partial specific volumes of the components, rCpG and ethidium were obtained by sedimentation on the isolated species.

Integration and rearrangement of equation (7) gives a linear equation where a plot of ln(C) versus r^2 has a slope from which $M_{complex}$ can be calculted. Table XIV shows the conditions for the equilibrium sedimentation and gives the measured partial specific volumes of the component species. The molecular weight of the complex was measured at a concentration where all of the ethidium was bound in a 2:1 complex, and under similar conditions to those used in the

TABLE XIV

CONDITIONS AND RESULTS OF EQUILIBRIUM SEDIMENTATION

Samples were run in standard buffer with

ρ = 1.0094 gm/cc

$\bar{V}_{NaCl}$ = 0.3 cc/gm

T = 278°K

speed = 42000 rpm

Partial Specific Volumes

$rCpG^{-}Na^{+}$ 0.58 cc/gm 2.5×10^{-5} M
measured from sedimentation run on rCpG solution

$Et.^{+}Cl^{-}$ 0.75 cc/gm
measured from sedimentation run on 0.1 mM Ethidium solution

2:1rCpG:Et complex 0.62 cc/gm
calculated from equation (8)

$M_{complex}$ = 1,800 gm/mole

calculated using equation (7) on solution of 2.5 mM rCpG 0.1 mM ethidium.

rest of the optical studies. Figure 4.5 shows the plot of ln(C) versus r^2 for the rCpG ethidium complex. The measured molecular weight of the complex was 1,800 gm/mole. This result shows that aggregates do not exist in solution to any appreciable extent since the molecular weight of a 2:1 complex would be 1,600, in good agreement with the measured value. Although the sedimentation equilibrium results cannot rule out a 2:2 complex, the earlier discussion on the stoichiometry showed that a 2:2 complex was not possible in our system.

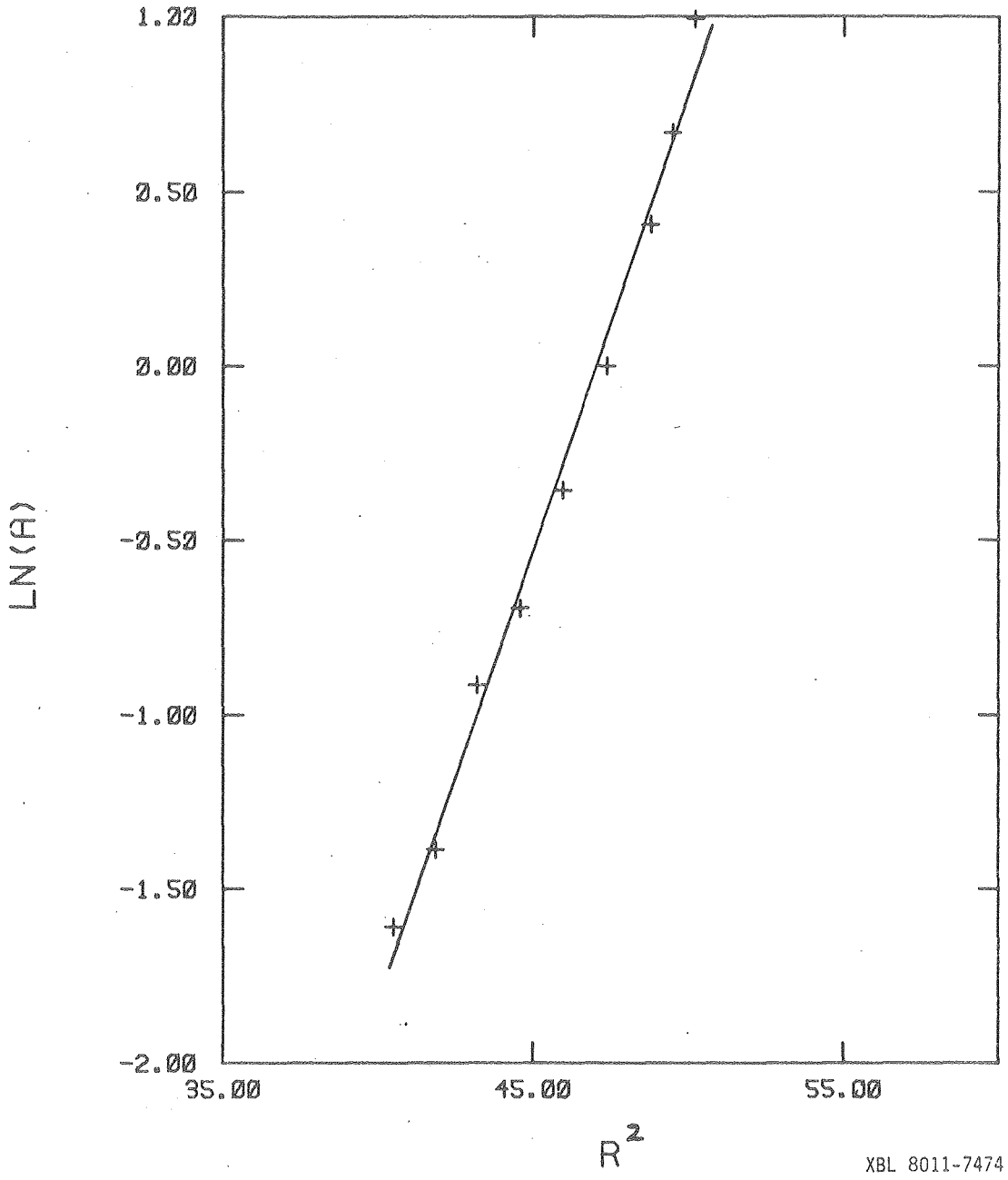
G) Induced CD of Ethidium upon Binding to Oligonucleotides

The induced CD of ethidium was measured when bound to the following oligonucleotides: rCpG, dCpG, and rUpA. The CD from 290 to 420 nm is shown in Figure 4.6 for these three dimers. The epsilon left - epsilon right ($\Delta\varepsilon_b$) shown in Figure 4.6 is per ethidium bound. The spectra were measured under the same conditions (and often on the same solutions) as those used in the equilibrium constant determinations. The $\Delta\varepsilon_b$ values were calculated from the relation

$$\theta^\circ = 32.98(\Delta A) = 32.98(C_{complex}\Delta\varepsilon_b \ell)$$

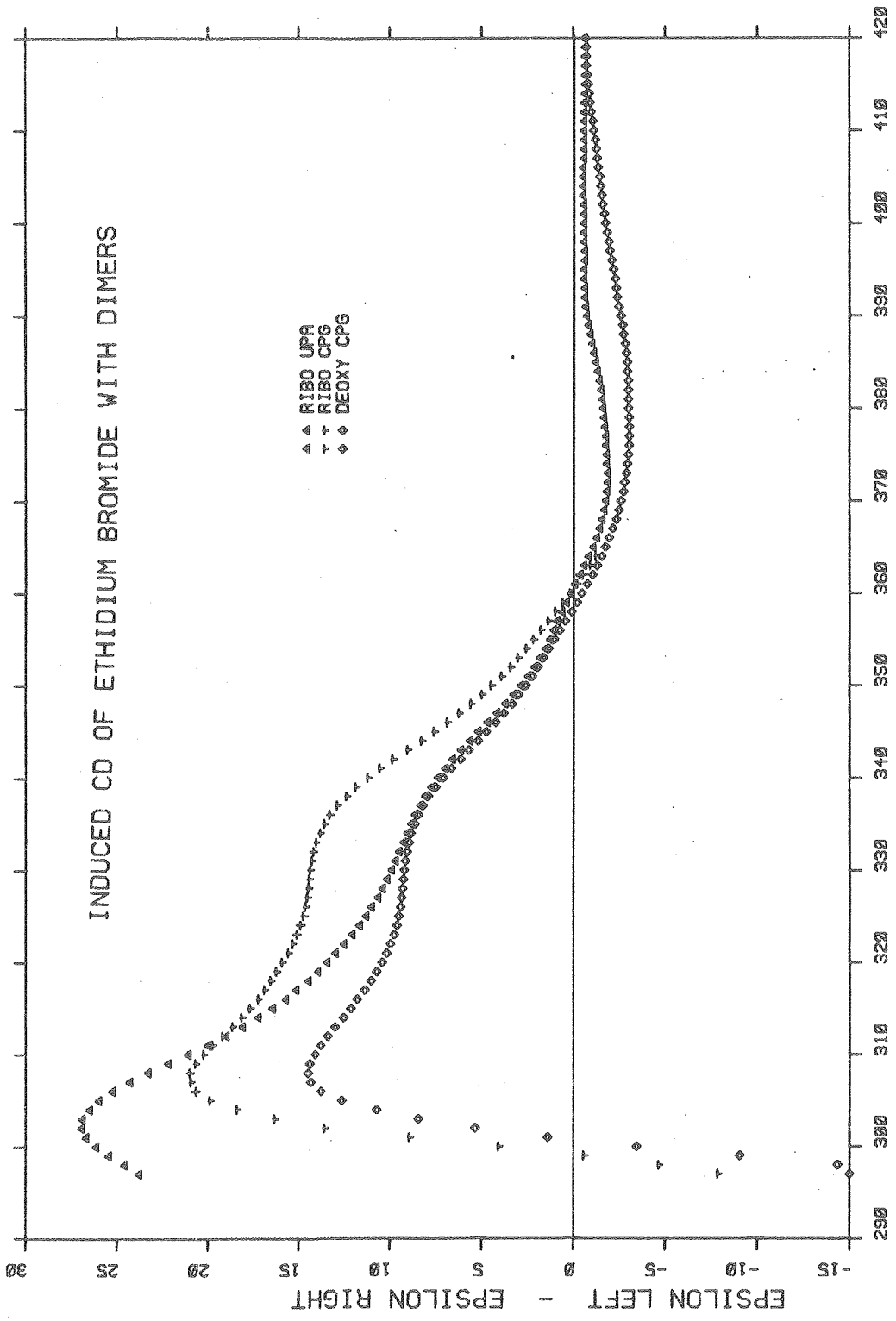
where θ° is the measured ellipticity in degrees, $C_{complex}$ is the concentration of complex (calculated using the K's given earlier) and ℓ is the pathlength of the cell in cm.

Figure 4.5. ln(A) versus r^2 plot for 2.5 mM rCpG and 0.1 mM ethidium. This is equivalent to ln(C) vs. r^2 since $A \propto C$.



XBL 8011-7474

Figure 4.6. Induced CD of ethidium when bound to rUpA, rCpG, or dCpG.


INDUCED CD OF ETHIDIUM BROMIDE WITH DIMERS
RIBO UPA
RIBO CPG
DEOXY CPG
EPSILON LEFT - EPSILON RIGHT
WAVELENGTH (NM)
30
25
20
15
10
5
0
-5
-10
-15
290
300
310
320
330
340
350
360
370
380
390
400
410
420

XBL 795-9565

The maximum in the CD of ethidium is around 307 nm, and the $\Delta\varepsilon_{b307's}$ for the three dimer-ethidium complexes are slightly different, with rUpA being the largest and dCpG the smallest (see Figure 4.6). For a given dimer the calculated $\Delta\varepsilon_{b307}$ is fairly constant throughout the titration curve as seen in Table XV. This will be important when comparing the induced CD obtained from ethidium-dimer complexes with those from ethidium-DNA complexes.

Figure 4.7 shows the induced CD of ethidium upon binding to oligonucleotides with a bulge on one strand (rGpUpG + rCpC), and with bulges on both strands (rCpUpG). This data was taken and kindly provided by Ms. Barbara Dengler. The conditions are slightly different than those used in this study (no buffer or salt was added, pH = 7.0, and the ratio was 2:1 nucleotides:ethidium) so the equilibrium constants obtained earlier are probably not valid here. The epsilon left - epsilon right in Figure 4.7 is not per ethidium bound and therefore not comparable with that in Figure 4.6. The shapes of the induced CD for rGpUpG + rCpC + Et. and rCpUpG + Et. are very similar to those of the dinucleosides with ethidium shown in Figure 4.6. This indicates that the property of the oligonucleotides which causes the induced CD is present in complexes with bulges as well as those with perfect helices.

TABLE XV

MAGNITUDE OF THE INDUCED CD OF ETHIDIUM AT 307 nm DURING TITRATION CURVE

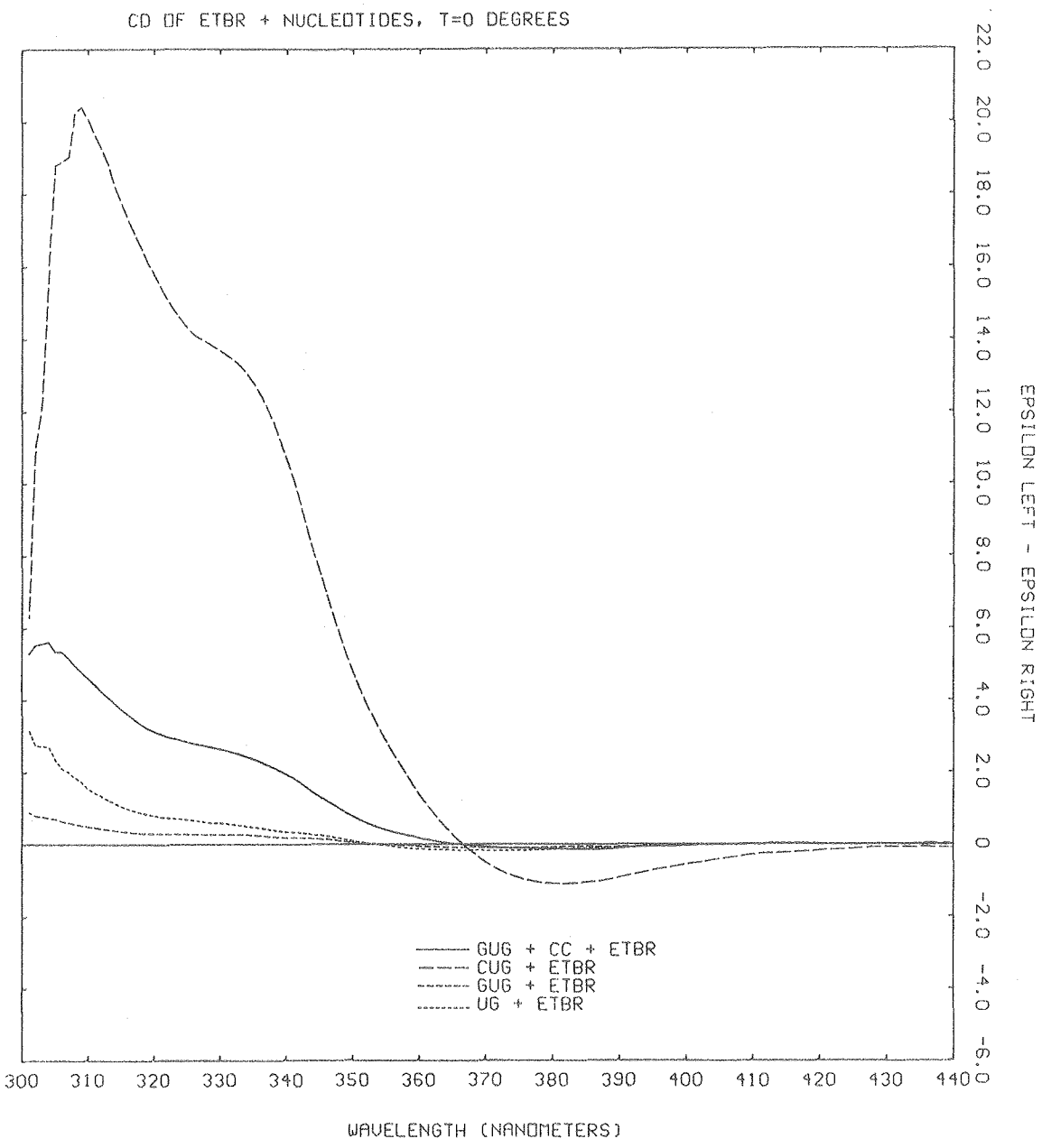
rCpA 0°C

[rCpG] mM	α	$\Delta\varepsilon_{b\ 307}$ $M^{-1}\ cm^{-1}$
0.13	0.47	28
0.20	0.67	24
0.27	0.80	24
0.4	0.91	22.5
1.0	0.99	23.5
2.0	1.00	23.3

dCpG 0°C

[dCpG] mM	α	$\Delta\varepsilon_{b307}$ $M^{-1}cm^{-1}$
0.35	0.49	12.6
0.5	0.67	13.2
0.70	0.79	11.5
1.0	0.89	15.2
2.0	0.98	14.4
3.0	0.99	11.8

Figure 4.7. Induced CD of ethidium when bound to rCpUpG, rGpUpG + rCpC, rGpUpG, or rUpG.

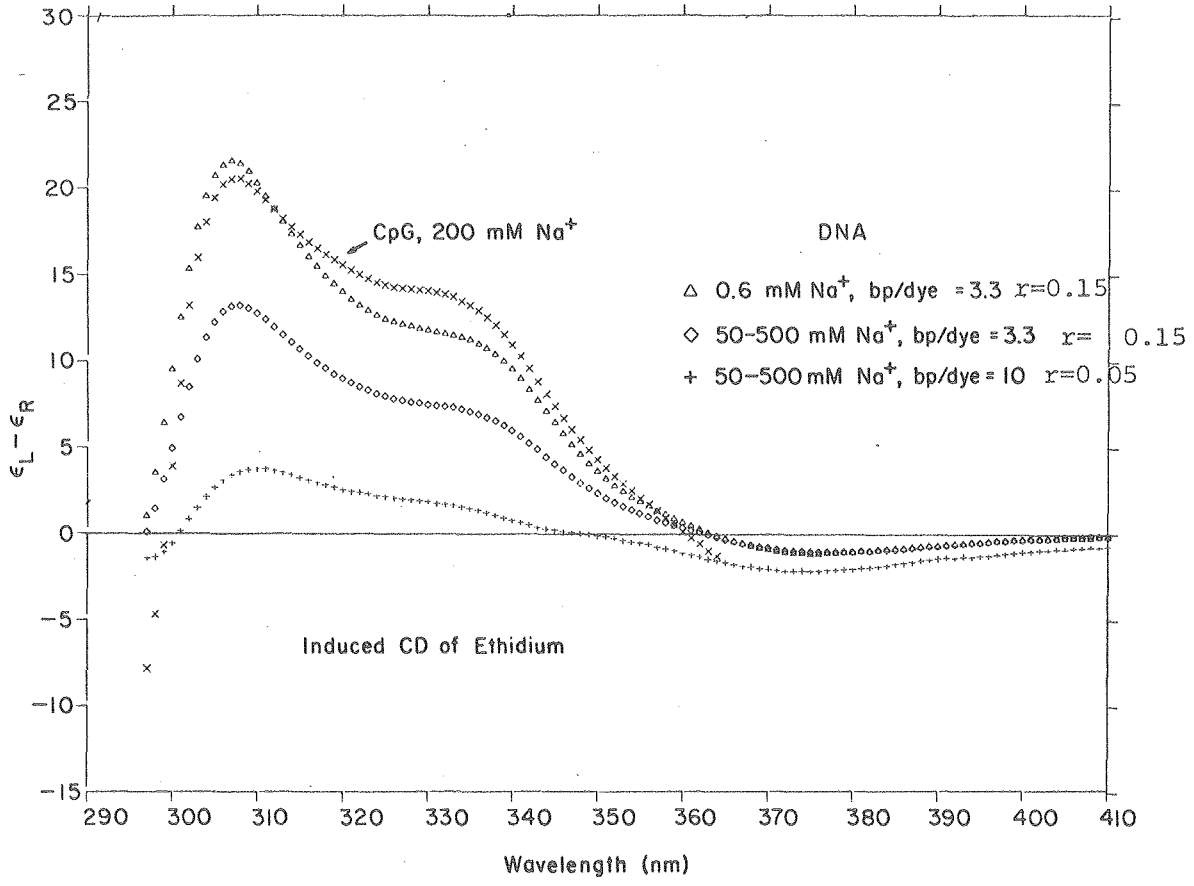

CD OF ETBR + NUCLEOTIDES, T=0 DEGREES
22.0 20.0 18.0 16.0 14.0 12.0 10.0 8.0 6.0 4.0 2.0 0 -2.0 -4.0 -6.0
EPSILON LEFT - EPSILON RIGHT
GUG + CC + ETBR
CUG + ETBR
GUG + ETBR
UG + ETBR
300 310 320 330 340 350 360 370 380 390 400 410 420 430 440
WAVELENGTH (NANOMETERS)

XBL 784-8169

H) Induced CD of Ethidium upon Binding to *E. coli* DNA

The CD of ethidium upon binding to DNA has been thoroughly studied in the past.[79,80] The magnitude of the CD per ethidium bound at 307 nm ($\Delta\varepsilon_{b307}$) changes with r, the amount of ethidium bound per phosphate, in the same manner for many different types of DNA.[80] The reason for this change is not known but proposed mechanisms will be discussed later. In order to study this phenomenon we tried to perturb the system in such a manner that, for a set value of r, we obtain a change in the $\Delta\varepsilon_{b307}$ for the ethidium. An experiment such as this was done by Aktipis and Kindelis but they found no effect on the change of the value of $\Delta\varepsilon_{b307}$ with r when the salt concentration was changed from 40 mM to 5 M NaCl.[81] Thus, higher salt concentrations do not seem to affect the behavior of the induced CD. Here we chose to lower the salt concentration to see how this affected the induced CD of ethidium when bound to DNA. Figure 4.8 shows the CD of ethidium when bound to DNA under different conditions, and also compares these with the spectrum of ethidium when bound to the dimer rCpG. The two curves for the 50 - 500 mM Na^+ salt range show the variation in $\Delta\varepsilon_{b307}$ with r. From comparing the 0.6 mM and 50 - 500 mM Na^+ curves with the same values of r, we see that the magnitude of $\Delta\varepsilon_{b307}$ is much larger for the low salt curve. This means that by changing the amount of salt in solution, we have changed the value of

Figure 4.8. Induced CD of ethidium binding to rCpG or *E. coli* DNA. The amount of ethidium bound per phosphate, r, was obtained by looking at the visible absorption of the ethidium and from knowing the total amount of DNA in the sample. The ethidium was essentially all bound in these samples with DNA.


30
25
20
15
10
5
0
−5
−10
−15
290
300
310
320
330
340
350
360
370
380
390
400
410
$\epsilon_L - \epsilon_R$
Wavelength (nm)
CpG, 200 mM Na^+
DNA
△ 0.6 mM Na^+, bp/dye = 3.3 r=0.15
◇ 50-500 mM Na^+, bp/dye = 3.3 r= 0.15
+ 50-500 mM Na^+, bp/dye = 10 r=0.05
Induced CD of Ethidium

XBL 793-8897

$\Delta\varepsilon_{b307}$ without changing r. Therefore the amount of ethidium bound per phosphate (r) is not the only important parameter in the magnitude of the induced CD of ethidium when bound to DNA.

Comparison of the CD spectra of E. coli DNA without ethidium in 50 mM and 0.6 mM sodium ion, shows changes in the magnitude and shapes of the two curves. There was approximately a 15% decrease in the measured ellipticity at 257 nm in going from the 50 mM to 0.6 mM sodium ion. This difference is probably due to changes in the conformation of the DNA, which may be an important factor in the behavior of the induced CD of ethidium when bound to DNA.

4. DISCUSSION AND CONCLUSIONS

A) Thermodynamics of Ethidium Binding to Oligonucleotides

Ethidium has been studied in complexes with dinucleoside monophosphates,[72,77] and dinucleotide diphosphates,[82] by spectroscopic methods. These studies found that minihelices of two dimers per one ethidium were formed in solution. They also found that many of the properties of ethidium when bound in these dimer complexes were similar to those when ethidium intercalates in DNA. A sequence dependence for ethidium binding was found where ethidium prefers to bind to pyrimidine (3'-5') purine sites compared to purine (3'-5') pyrimidine sites.[72]

In this study we were able to measure the equilibrium constants and ΔH° values for ethidium binding in a complex which has a stoichiometry of two oligonucleotides and one ethidium. The equilibrium constants at 0°C were measured with the strength of the complex being

$$rCpG > dCpG > rCpUpG \approx rUpA > rGpUpG + rCpC.$$

There is a sequence dependence of the binding, with the rCpG complex having a much higher equilibrium constant than the rUpA. At least part of this difference is due to the lower stability of the A·U base pairs compared to the C·G base pairs. We also saw an order of magnitude higher K for binding of ethidium to the ribose dimer CpG compared to the deoxyribose dimer CpG. This difference may reflect the lower stability of deoxy C·G base pairs compared to ribose C·G base pairs.

A complex of ethidium with oligonucleotides which can form a bulge on one strand (rGpUpG + rCpC) was found to have a much smaller equilibrium constant than the rCpUpG-ethidium complex, which can form bulges on both strands. The rCpUpG complex is still much less stable than the perfect helix rCpG. All these oligonucleotides are too unstable to measure their equilibrium constants in the

absence of ethidium. Thus it is not possible to tell if the differences in the strength of the ethidium-oligonucleotide complexes are due to differences in the stability of the oligonucleotide helices themselves, or the strength of ethidium binding to that particular helix.

B) Induced Circular Dichroism of Ethidium

Ethidium is not optically active and hence has no intrinsic CD. A large positive CD from 300 - 360 nm for ethidium has been observed upon binding to DNA.[79] For this wavelength range the magnitude of the CD per ethidium bound has been shown to increase dramatically with the amount of ethidium bound.[80,81] A maximum is reached at an r value of 0.25, which is the neighbor exclusion limit for dye binding to DNA. This means that the magnitude of the induced CD for an individual ethidium molecule is dependent on how many other ethidiums are also bound to the nucleic acid. Several mechanisms have been proposed for this induced CD of ethidium upon binding to the DNA.[79,73] One proposal involves a direct interaction between two or more ethidium molecules bound in neighboring sites. Exciton interactions between the transition moments of these ethidiums bound in close proximity, will give rise to a conservative CD for this band. The change in $\Delta\varepsilon_{b307}$ with r is just a reflection of the fact that more dye molecules are closer together at higher r values.

The second model states that increasing number of ethidium molecules alter the environment of other dye molecules bound nearby, by changing the local DNA conformation. Thus the change in the $\Delta\varepsilon_{b307}$ for ethidium with r would be due to changes in the conformation of the ethidium-DNA complex. For this model the induced CD could arise from the assymetry of the binding site, or from the exciton interactions between the ethidium and the aromatic base transition moments.[79]

To get a better understanding of the induced CD of ethidium we have studied its binding to oligonucleotides and to DNA (as seen in Figures 4.6 - 4.8). The $\Delta\varepsilon_{b307}$ for the rCpG complex is 24±2.5 $M^{-1}cm^{-1}$, whereas Aktipis and Kindelis found this value to be 25, for calf thymus DNA at an r value of 0.25.[83] The earlier discussion on the dimer-ethidium complexes showed that they existed in solution as complexes of two dimers to one ethidium. The equilibrium sedimentation studies showed that there was no aggregation of these complexes, so the induced CD with the dimers arises from only one ethidium interacting with two dimers. The induced CD of the ethidium upon binding to dimers or DNA therefore cannot originate from dye-dye interactions.

As discussed earlier, the induced CD of ethidium may arise from base-dye interactions. Evidence against this mechanism comes from the similar induced CD of ethidium in

complexes of rUpA and rCpG. If the major cause of the CD is from the coupling of the transition dipoles of the ethidium and the bases, then we might expect quite different spectra for these two dimers which have no bases in common.

The induced CD may simply be a function of the conformation of the DNA around the ethidium binding site. We cannot rule out this possibility, by making this hypothesis our results indicate that the conformation of the ethidium-dimer complexes would be similar to the conformation of the ethidium when it is bound to DNA near the neighbor exclusion limit.

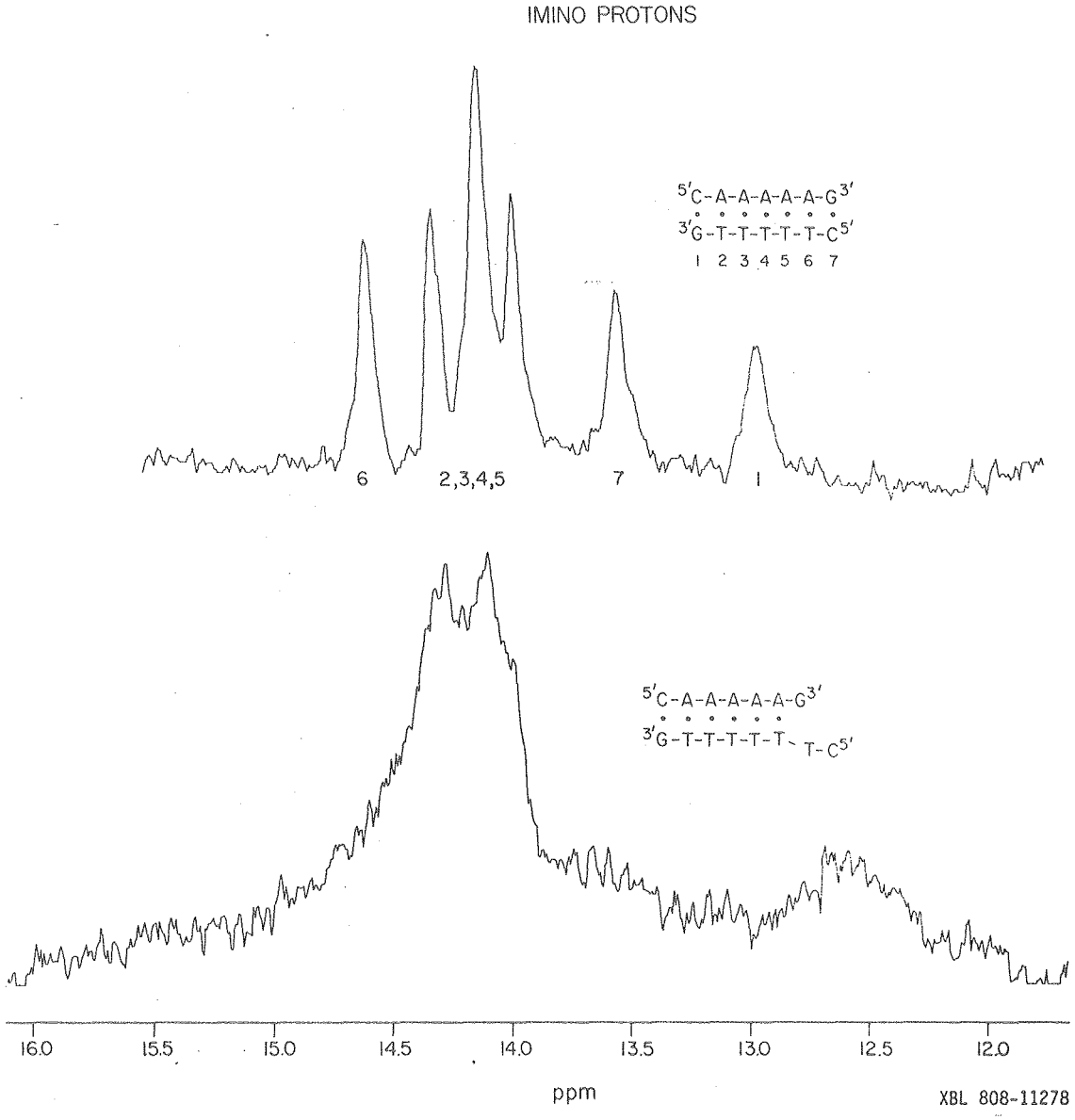
The interaction of the amino groups on the ethidium with the phosphates in the nucleic acids, may be an important contribution to the induced CD of the ethidium. Molecular orbital calculations on ethidium have shown that the major contribution to the large red shift in the visible absorbance of ethidium is the interaction of the ethidium with the charged phosphates.[84] The induced CD of ethidium upon binding to E. coli DNA changed by going to very low salt (where the charge on the phosphates may be different than at higher salt levels).[85] This indicates interaction of ethidium with the charged phosphates may be important in the induced CD of ethidium when bound to nucleic acids. These questions and others are presently being pursued by Mr. Kenneth Dahl, who I wish to thank for providing valuable advice and help on this project.

APPENDIX I

NMR STUDY OF A DOUBLE HELIX WITH A MISMATCHED BASE

The imino proton region of the helix dCA_5G + dCT_6G was observed from 1 - 15°C. These molecules can form a double helix with a bulged thymine on one strand. If the C·G base paris on both strands are formed, then we would expect to see their imino protons in the NMR spectrum. What was observed for this helix was a broad peak for all the A·T imino protons, and another very broad, much smaller, peak at higher field presumably due to C·G protons. This is illustrated in Figure I.1, which gives the spectra for the perfect helix dCA_5G + dCT_5G and the mismatched helix dCA_5G + dCT_6G. The spectrum of the mismatched helix shows that both C·G base protons are not formed, and is indicative of dangling ends on this helix (the structure of a dangling end is shown above the spectrum in Figure I.1). Since the bulge did not seem to form, even at 1°C, we tried to stabilize that structure by changing the conditions in solution. The sample was run in a mixture of 1:2 methanol:H_2O (v/v) at -20°C. Essentially no change in the spectrum was observed. Other conditions attempted were, 1 M NaCl, 1°C; 1 M NaCl + 5 mM Mg^{2+} in H_2O, and 1 mM spermidine in our standard buffer at 1°C. In none of these experiments were we able to observe a spectrum which had two C·G imino protons, which indicates

Figure I.1 Comparison of the imino proton region of a perfect helix, $dCT_5G + dCA_5G$, and a helix which has a mismatched base on one strand, $(dCA_5G + dCT_6G)$.


IMINO PROTONS
5′C-A-A-A-A-A-G3′
3′G-T-T-T-T-T-C5′
1 2 3 4 5 6 7
6
2,3,4,5
7
1
5′C-A-A-A-A-A-G3′
3′G-T-T-T-T-T- T-C5′
16.0
15.5
15.0
14.5
14.0
13.5
13.0
12.5
12.0
ppm
XBL 808-11278

that we were not getting a helix with a bulged thymine. The bulged structure must be energetically less favorable than the dangling end in this helix. Studies on helices which can form a bulge, but not a dangling end, will be attempted in the future.

APPENDIX II

COMPUTER PROGRAMS

Computer programs were written in Fortran on the Vax system at the Laboratory of Chemical Biodynamics.

WIDTH2 calculates the lineshape of a two state system with exchange between the two states. The input parameters are the chemical shifts, the linewidths, and the lifetimes of the two states. The program will then calculate the chemical shift of the largest peak and its linewidth. It will also plot the whole spectrum using a program called PLOT (Mr. David Keller).

ITWIDT is a program similar to WIDTH2, but it also finds the best fit for the experimental and calculated chemical shift for a two state system. It does this by varying the mean exchange lifetime of the system, for a given population, until the best fit between the calculated and experimental results are obtained.

ITWIDT

```
c       This program is similar to the WIDTH2 program and the documentation
c       for WIDTH2 explains most of the parameters needed here.  The
c       difference is that ITWIDT will vary the mean lifetime (tau) of
c       the two state system until the best fit between the experimental
c       and calculated chemical shift are found.  It does this for
c       a given population and where all the other parameters are fixed.
c       Thus it will increment the population of the two states, and
c       find the lifetime which gives the best fit between the experimental
c       and calculated chemical shift of the largest peak in the spectrum.
c
c
c
c
c
c
      COMMON V,DELTA,DELTAA,DELTAB,T2A,T2B,POPA,POPB,DMAX,TMAX,PI
      DIMENSION V(5000),TTAU(400)
      DOUBLE PRECISION DELTA(5000),TAU
111   CONTINUE
      TYPE *,'THE TEMPERATURE IS =  (I3)'
      READ 110, ITEMP
      TYPE *,'THE MINIMUM POPULATION OF STATE A IS      POPAMIN=  (F5.2)'
      READ 200,POPAMIN
      TYPE *,'THE MAXIMUM POPULATION OF STATE A IS     POPAMAX=   (F5.2)'
      READ 200, POPAMAX
      TYPE *,'THE INTERVAL FOR CALCULATING THE POPULATION = (F5.2)'
      READ 200, THEIN
      TYPE *,'THE CHEMICAL SHIFT OF A IS    DELTAA=  (F10.3) IN HZ'
      READ 300,DELTAA
      TYPE *,'THE CHEMICAL SHIFT OF B IS    DELTAB=  (F10.3) IN HZ'
      READ 300,DELTAB
      TYPE *,'THE EXPERIMENTAL CHEMICAL SHIFT IS = (F10.3) IN HZ'
      READ 300,EXMAX
      TYPE *,'LINEWIDTH OF STATE A IS EQUAL TO     (E12.5)'
      READ 100,T2A
      TYPE *,'LINEWIDTH OF STATE B IS EQUAL TO     (E12.5)'
      READ 100,T2B
      PI=3.141592
      T2=T2*PI
      DELTA(1)=(DELTAA + DELTAB)/2. - 200.
      ID=INT(DELTA(1))
      DELTA(1)=FLOAT(ID)
      DO 50 I=2.4000
      DELTA(I)=DELTA(I-1) + 0.1
50    CONTINUE
      POPA=POPAMIN - THEIN
      WRITE(2,510)ITEMP
      WRITE(2,700)DELTAA,DELTAB
      WRITE(2,710)EXMAX
      WRITE(2,800)T2A/PI
      WRITE(2,801)T2B/PI
      WRITE(2,860)
      DO 1100 J=1,(POPAMAX-POPAMIN)/THEIN +1
      POPA=POPA + THEIN
      POPB=1.-POPA
      TAUMAX=ABS(40./(2.*PI*(DELTAA - DELTAB)))
      TAUMIN=ABS(0.05/(2.*PI*(DELTAA - DELTAB)))
1150  CONTINUE
      SIN=ABS((LOG10(TAUMAX)-LOG10(TAUMIN))/20.)
      IF (TAUMAX .GT. TAUMIN ) THEN
          XY=TAUMIN
      ELSE
```

```
          XY=TAUMAX
        END IF
        TTAU(1)=LOG10(XY)
        DMAX1=1000.
        DMAX2=2000.
        TAU1=XY
        DO 1111 M=1,20
        TTAU(M +1)=TTAU(M) + SIN
        TAU=10.**TTAU(M)
        CALL GETMAX(TAU)
        X=ABS(DMAX-EXMAX)
        IF (X .LT. DMAX1) THEN
          DMAX2=DMAX1
          DMAX1=X
          TAU2=TAU1
          TAU1=TAU
          GO TO 1111
        END IF
        IF(X .LT. DMAX2) THEN
          DMAX2=X
          TAU2=TAU
        END IF
1111    CONTINUE
        X2=ABS(TAU1-TAU2)
        X3=ABS(0.01*TAU1)
        IF (X3 .GT.X2) GO TO 1090
1110    CONTINUE
        TAUMAX=TAU1
        TAUMIN=TAU2
        GO TO 1150
1090    CONTINUE
        TAU=TAU1
        CS=POPA*T2B +POPB*T2A
        DO 60 I=1,4000
        SP=TAU*(T2A*T2B - (PI*(DELTAA + DELTAB) -2.*PI*DELTA(I))**2 +
     2  (PI*(DELTAA - DELTAB))**2) + POPA*T2A + POPB*T2B
        SQ=TAU*(PI*(DELTAA + DELTAB) - 2.*PI*DELTA(I) - (POPA -
     2   POPB)*PI*(DELTAA - DELTAB))
        SR=(PI*(DELTAA+DELTAB)-2.*PI*DELTA(I))*(1.0+TAU*(T2B+T2A))
     2     + (POPA -POPB)*(DELTAA - DELTAB)*PI +
     3  TAU*PI*(DELTAA-DELTAB)*(T2B-T2A)
        V(I)=(1. + TAU*CS)*  SP + SQ*SR)/(SP**2 + SR*SR)
60      CONTINUE
        V(1)=0.
        DO 190 I=2,4000
        IF (V(I) .GT. TMAX) THEN
           TMAX=V(I)
           DMAX=DELTA(I)
        END IF
190     CONTINUE
        SNEARZ=1000.
        DO 210 I=1,4000
        IF (ABS(TMAX/2.-V(I)) .LT. SNEARZ) THEN
           HMAX=DELTA(I)
           SNEARZ=TMAX/2-V(I)
        END IF
210     CONTINUE
        EXCH=ABS(2.*PI*(DELTAA-DELTAB)*TAU)
        WIDTH=ABS((DMAX-HMAX)*2.)
        POPAMX=(DMAX -DELTAB)/(DELTAA-DELTAB)
        POPBMX= 1. -POPAMX
        WRITE(2,500)POPA,TAU,DMAX,WIDTH,POPAMX,EXCH
1100    CONTINUE
```

```
 710     FORMAT('  THE EXPERIMENTAL CHEMICAL SHIFT IS = 'F10.3)
 800     FORMAT('  INTRINSIC LINEWIDTH OF PEAK A IS EQUAL TO ='E12.5)
 801     FORMAT('  INTRINSIC LINEWIDTH OF PEAK B IS EQUAL TO ='E12.5)
100      FORMAT(E12.5)
110      FORMAT(I3)
112      FORMAT(////////////)
200      FORMAT(F5.2)
300      FORMAT(F10.3)
500      FORMAT(10X,F5.2,T30,E10.2,T45,F7.1,13X,F7.1,18X,F5.2,T110,F8.2)
510      FORMAT('  THE TEMPERATRUE OF THIS  DATA SET IS 'I3)
700      FORMAT('  THE CHEMICAL SHIFT OF STATE A IS = ',F8.2,/
     2   '   THE CHEMICAL SHIFT OF STATE B IS = ',F8.2)
860      FORMAT(///,T10,'INPUT POPULATION',T30,'CALCULATED ',T45,
     2   'CALCUALTED',T65,'CALCULATED WIDTH',T85,'CACULATED POPULATION',
     3   T115,'EXCHANGE ',/,T12,'OF STATE A ',T35,'TAU',T50,
     4   'MAX',T88.'OF STATE A',T115,'PARAMETER'/)
         WRITE(2,112)
         TYPE *,'DO YOU WANT TO ENTER ANOTHER DATA SET? 1=YES,2=NO'
         READ 110,IN
         IF (IN .EQ. 1) GO TO 111
         END
C
C
C
C
C
         SUBROUTINE GETMAX(TAU)
         DIMENSION V(5000)
         DOUBLE PRECISION DELTA(5000)
         COMMON V,DELTA,DELTAA,DELTAB,T2A,T2B,POPA,POPB,DMAX,TMAX,PI
         CS=POPA*T2B +POPB*T2A
         DO 60 I=1,4000,10
         SP=TAU*(T2A*T2B - (PI*(DELTAA + DELTAB) -2.*PI*DELTA(I))**2 +
     2   (PI*(DELTAA - DELTAB))**2) + POPA*T2A + POPB*T2B
         SQ=TAU*(PI*(DELTAA + DELTAB) - 2.*PI*DELTA(I) - (POPA -
     2    POPB)*PI*(DELTAA - DELTAB))
         SR=(PI*(DELTAA + DELTAB)-2.*PI*DELTA(I))*(1.0+TAU*(T2B+T2A))
     2    + (POPA -POPB)*(DELTAA - DELTAB)*PI + TAU*PI*(DELTAA-DELTAB)
     3   * (T2B-T2A)
         V(I)=((1. + TAU*CS)*SP + SQ*SR)/(SP**2 + SR*SR)
60       CONTINUE
         TMAX=0.0
         DO 190 I=1,4000,10
         IF (V(I) .GT. TMAX) THEN
            TMAX=V(I)
            DMAX=DELTA(I)
         END IF
 190     CONTINUE
         RETURN
         END
```

WIDTH2

```
C       This program calculates the line shape of nmr lines when there is
C       chemical exchange between two states.  The equation and theory
C       used are from HIGH RESOLUTION NUCLEAR MAGNETIC RESONANCE by Pople
C       Schneider, and Berstein, page 222.
C
C
C
C               The parameters which you must enter are the chemical shifts
C       of the two states, the population of the two states, the average
C       T2 of the two states(this derivation assumes that both T2's are
C       the same), and the lifetimes of the two states.
C               It then calculates the lineshapes of the peaks for an area
C       which is plus or minus 200 Hz of the average chemical shifts of
C       the two resonances.  Therefore if the chemical shifts of the two
C       states differ by much more than 350 Hz or so, this program should
C       be modified to account for this large area.  The program outputs
C       a file by the name FOR002.DAT;(and some file extension) which has
C       the intensity of the peak over the chemical shift range designated
C       One can then set up a plot file and plot the data if so desired.
C
C
C
C
C       The plot program is called PLOT
        DIMENSION V(5000)
        DOUBLE PRECISION DELTA(5000)
        TYPE *,'THE TEMPERATURE IS =  (I3)'
        READ 110, ITEMP
        TYPE *,'THE VALUE OF tauA IS =   (E12.5)'
        READ 100,TAUA
        TYPE *,'THE VALUE OF TAUB IS =   (E12.5)'
        READ 100,TAUB
        TYPE *,'THE POPULATION OF STATE A IS      POPA=  (F5.2)'
        READ 200,POPA
        TYPE *,'THE CHEMICAL SHIFT OF A IS    DELTAA=  (F10.3) IN HZ'
        READ 300,DELTAA
        TYPE *,'THE CHEMICAL SHIFT OF B IS    DELTAB=  (F10.3) IN HZ'
        READ 300,DELTAB
        TYPE *,'INTRINSIC LINEWIDTH OF PEAK A  IS EQUAL TO    (E12.5)'
        READ 100,T2A
        TYPE *,'INTRINSIC LINEWIDTH OF PEAK B IS EQUAL TO   (E12.5)'
        READ 100,T2B
        TAU=TAUA*TAUB/(TAUA+TAUB)
        PI=3.141592
        POPB=1.-POPA
        SUME=0.0
        T2A=PI*T2A
        T2B=PI*T2B
        DELTA(1)=(DELTAA + DELTAB)/2. - 200.
        ID=INT(DELTA(1))
        DELTA(1)=FLOAT(ID)
        DO 50 I=2,4000
        DELTA(I)=DELTA(I-1) + 0.10000
50      CONTINUE
        CS=POPA*T2B+POPB*T2A
        SUM=0.
        DO 60 I=1.4000
        SP=TAU*(T2A*T2B - (PI*(DELTAA + DELTAB) -2.*PI*DELTA(I))**2 +
     2  (PI*(DELTAA - DELTAB))**2) + POPA*T2A +POPB*T2B
        SQ=TAU*(PI*(DELTAA + DELTAB) - 2.*PI*DELTA(I) - (POPA -
     2   POPB)*PI*(DELTAA - DELTAB))
        SR=(PI*(DELTAA + DELTAB) - 2.*PI*DELTA(I))* (1.0 + TAU*(T2A+T2B))
```

```
     2    + (POPA -POPB)*(DELTAA - DELTAB)*PI + TAU*PI*(DELTAA-DELTAB)*
     3   (T2B-T2A)
         V(I)=((1. + TAU*CS)*SP + SQ*SR)/(SP**2 + SR*SR)
60       CONTINUE
         DO 70 I=1,3999
         SUM=SUM + (V(I) + V(I+1))/2.*0.1
70       CONTINUE
         TK=SUM
         DO 80 I=1.4000
         V(I)=V(I)*100./SUM
80       CONTINUE
         WRITE(2,600) (DELTA(I),V(I), I=1,4000,10)
         WRITE(9,601) (DELTA(I),V(I), I=1,4000,10)
         DO 90 I=1,3999
         SUME=SUME + (V(I)+V(I+1))/2.*0.1
90       CONTINUE
         V(1)=0.
         DO 190 I=2,4000
         IF (V(I) .GT. TMAX) THEN
            TMAX=V(I)
            DMAX=DELTA(I)
            IM=I
         END IF
190      CONTINUE
         SNEARZ=1000.
         DO 210 I=IM,IM+500
         IF (ABS(TMAX/2.-V(I)) .LT. SNEARZ) THEN
            HMAX=DELTA(I)
            SNEARZ=ABS(TMAX/2-V(I))
         END IF
         IF (SNEARZ .LT. 0.01*TMAX) GO TO 2100
210      CONTINUE
2100     CONTINUE
         WIDTH= ABS((DMAX-HMAX)*2.)
         POPAMX=(DMAX - DELTAB)/(DELTAA-DELTAB)
         POPBMX= 1. -POPAMX
         WRITE(2,510)ITEMP
         WRITE(2,400)TAUA,TAUB
         WRITE(2,500)POPA,POPB
         WRITE(2,700)DELTAA,DELTAB
         WRITE(2,660)WIDTH, DMAX
         WRITE(2,670)POPAMX,POPBMX
         WRITE(2,800)T2A/PI
         WRITE(2,801)T2B/PI
         WRITE(9,510)ITEMP
         WRITE(9,400)TAUA,TAUB
         WRITE(9,500)POPA,POPB
         WRITE(9,700)DELTAA,DELTAB
         WRITE(9,660)WIDTH,DMAX
         WRITE(9,670)POPAMX,POPBMX
         WRITE(9,800)T2A/PI
         WRITE(9,801)T2B/PI
400      FORMAT(//'  THE ABOVE DATA IS FOR THE FOLLOWING PARAMETERS'///
     2   '  The value of tauA is = 'E12.5/'  The value of tauB is = '
     3   E12.5)
500      FORMAT(//'  The population of state A is ='F5.2/
     2   '  The population of state B is ='F5.2)
700      FORMAT(//'  The chemical shift of A is ='F10.3' Hz'/
     2   '  The chemical shift of B is ='F10.3' Hz')
800      FORMAT(' THE INTRINSIC LINEWIDTH OF STATE A IS EQUAL TO ='E12.5)
801      FORMAT(' THE INTRINSIC LINEWIDTH OF STATE B IS EQUAL TO ='E12.5)
601      FORMAT(1X,F10.5,F13.5)
600      FORMAT(410(6(1X,F6.1,3X,F5.2,4X)./))
```

```
100     FORMAT(E12.5)
110     FORMAT(I3)
200     FORMAT(F5.2)
300     FORMAT(F10.3)
660     FORMAT(///' The width at half max of the largest peak is  'F6.2
     2  'Hz'/' The chemical shift of the maximum of this peak is  '
     3  F6.1'Hz' )
670     FORMAT(////' If one assumes we are in fast exhcange and uses this
     2  chemical shift to calcualte the populations you get:'
     3  /' The calculated population of state A is =  'F5.2
     4  /' The calculated population of state B is =  'F5.2)
510     FORMAT(' THE TEMPERATRUE OF THIS  DATA SET IS 'I3)
        END
```

REFERENCES

1. Watson, J.D. and Crick, F.H.C. (1953) Nature 171, 964.

2. Arnott, S., Fuller, W., Hodgson, A. and Prutton, I. (1968) Nature 220, 561.

3. Arnott, S., Chandrasekaran, R. and Selsing, E. (1975), in Structure and Conformations of Nucleic Acids and Protein-Nucleic Acid Interactions (Sundaralingam, M. ED.) pp 577-796, University Park Press, Baltimore, MD.

4. Bloomfield, V.A., Crothers, D.M. and Tinoco, I., JR. (1974) Physical Chemistry of Nucleic Acids, pp 339-342, Harper and Row, New York.

5. Giessner-Prettre, C. and Pullman, B. (1970) J. Theor. Biol. 27, 87.

6. Robillard, G.T. and Reid, B.R. (1979) in Biological Applications of Magnetic Resonance (Shulman, R.G., Ed.) pp 45-112, Academic Press, New York.

7. Kearns, D.R. (1977) in Annual Review of Biophysics and Bioengineering (Mullins, L.F., Hagins, W.A., Stryer, L. and Newton, C. eds.), Vol. 6, pp 477-503, Academic Press, London.

8. Lee, C.H., Ezra, F.S., Kondo, N.S., Sarma, R.H. and Danyluk, S.S. (1976), Biochemistry 15, 3627.

9. Ezra, F.A., Lee, C.H., Kondo, N.S., Danyluk, S.S. and Sarma, R.H. (1977) Biochemistry 16, 1977.

10. Karplus, M. (1959) J. Chem. Phys. 30, 11.

11. McCann, J., Choi, E., Yamasaki, E. and Ames, B.N. (1975) Proc. Nat. Acad. Sci. 72, 5135.

12. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) Cold Spring Harbor Symposium on Quantitative Biology 31, 77.

13. Arter, D.B., Walker, G.C., Uhlenbeck, O.C. and Schmidt, P.G. (1974) Biochem. Biophys. Res. Commun. 61, 1089.

14. Heller, M.J., Anthony, T.T., and Maciel, G.E. (1974) Biochemistry 13, 1623.

15. Borer, P.N., Kan, L.S., and Ts'o, P.O.P. (1975) Biochemistry 14, 4847.

16. Hughes, D.W., Bell, R.A., England, T.E. and Neilson, T. (1978) Can J. Chem. 56, 2243.

17. Cross, A.D. and Crothers, D.M. (1971) Biochemistry 10, 4015.

18. Patel, D.J. (1974) Biochemistry 13, 2396.

19. Kallenbach, N.R., Daniel, W.E., Jr. and Kaminker, M.A. (1976) Biochemistry 15, 1218.

20. Early, T.A., Kearns, D.R., Burd, J.F., Larson, J.E. and Wells, R.D. (1977) Biochemistry 16, 541.

21. Milman, G., Langridge, R., and Chamberlin, M.J. (1967) Proc. Natl. Acad. Sci. 57, 1804.

22. Selsing, E., Wells, R.D., Early, T.A. and Kearns, D.R. (1978) Nature 21, 249.

23. Geerdes, H.A.M. and Hilbers, C.W. (1977) Nucl. Acids Res. 4, 207.

24. Martin, F.H. and Tinoco, I., JR. (1980) Nucl. Acids Res. 8, 2295.

25. Martin, F.H., Uhlenbeck, O.C. and Doty, P. (1971) J. Mol. Biol. 57, 201.

26. Patel, D.J. (1979) Eur. J. Biochem. 96, 267.

27. Patel, D.J. (1975) Biochemistry 14, 3984.

28. Khorang, H.G. (1968) Pure Appl. Chem. 17, 349.

29. Handbook of Biochemistry Selected Data for Molecular Biology 3rd Edition, CRC Press, pp 586.

30. Redifled, A.G., Kunz, S.D. and Ralph, E.K. (1975) J. Magn. Reson. 19, 114.

31. Ts'o, P.O.P., Barrett, J.C., Kan, L.S. and Miller, P.S. (1973) Ann. N.Y. Acad. Sci. 222, 290.

32. Cheng, D.M. and Sarma, R.H. (1977) J. Am. Chem. Soc. 99, 7333.

33. Chiao, Y.C.C. and Krugh, T.R. (1977) Biochem. 16, 747.

34. Dwek, R.A. (1973) Nuclear Magnetic Resonance in Biochemistry pp 213-326, London, Oxford University Press.

35. Barry, C.D., North, A.C.T., Glasel, J.A., Williams, R.J.P. and Xavier, A.V. (1971) Nature 232, 236.

36. Chan, S.I., Nelson, J.H., (1969) J. Am. Chem. Soc. 91, 168.

37. Lee, C.H., Tinoco, I., Jr. (1977) Biochem. 16, 5403.

38. Kearns, D.R. and Shulman, R.G. (1974) Acc. Chem. Res. 7, 33.

39. Sarma, R.H. (1980) Nucleic Acid and Geometry and Dynamics Pergamon Press, New York.

40. Kan, L.S., Borer, P.N. and Ts'o, P.O.P. (1975) Biochemistry 14, 4864.

41. Arter, D.B. and Schmidt, P.G. (1976) Nucl. Acids Res. 3, 1437.

42. Lee, C.H. and Tinoco, I., Jr. (1980) Biophys. Chem. 11, 283.

43. Ts'o, P.O.P., Kondo, N.S., Schweizer, M.P. and Hollis, D.P. (1969), Biochemistry 8, 997.

44. Nelson, J.W., Martin, F.H., and Tinoco, I., Jr., in preparation.

45. Gutowsky, H.S., and Holm, C.H. (1956), J. Chem. Phys. 25, 1228.

46. Kaplan, J.I. and Fraenkel, G. (1980) NMR of Chemically Exchanging Systems, pp 74-80, Academic Press, New York.

47. Altona, C., and Sundaralingam, M. (1973) J. Am. Chem. Soc. 95, 2333.

48. Porschke, D., and Eigen, M. (1971) J. Mol. Biol. 62, 361.

49. Ravetch, J., Gralla, J. and Crothers, D.M. (1973) Nucl. Acids Res. 1, 109.

50. Arnott, S. and Hukins, D.W.L. (1973) J. Mol. Biol. 81, 93.

51. Shum, B.W.K. (1977), Ph.D. Thesis, Yale University.

52. Feeney, J., Batchelor, J.G., Albrand, J.P. and Roberts, G.C.K. (1979) J. Magn. Reson. 33, 519.

53. Hilbers, C.W. (1979) in Biological Applications of Magnetic Resonance (Shulman, R.G., Ed.) pp 1-44, Academic Press, New York.

54. Patel, D.J. and Hilbers, C.W. (1975) Biochemistry 14, 2651.

55. Hilbers, C.W. and Patel, D.J. (1975) Biochemistry 14, 2657.

56. Crothers, D.M., Cole, P.E., Hibers, C.W. and Shulman, R.G. (1974) J. Mol. Biol. 87, 63.

57. Englander, J.J., Kallenbach, N.R. and Englander, S.W. (1972) J. Mol. Biol. 63, 153.

58. Teitelbaum, H. and Englander, S.W. (1975) J. Mol. Biol. 92, 55.

59. Teitelbaum, H. and Englander, S.W. (1975) J. Mol. Biol. 92, 79.

60. Kallenbach, N.R. (1980) in Nucleic Acid Geometry and Dynamics (Sarma, R.H. Ed.), pp 231-255, Pergamon Press, New York.

61. Waelder, S., Lee, L., and Redfield, A.G. (1975) J. Am. Chem. Soc. 97, 2927.

62. Johnston, P.D. and Redfield, A.G. (1977) Nucl. Acids Res. 4, 3599.

63. Johnston, P.D. and Redfield, A.G. (1978) Nucl. Acids Res. 5, 3913.

64. Shih, W.C.W. (1979) Ph.D. Thesis, University of California, Berkeley.

65. McConnell, H.M. (1958) J. Chem. Phys. 28, 430.

66. Campbell, I.D., Dobson, C.M. and Ratcliffe, R.G. (1977) J. Magn. Reson. 27, 455.

67. Marshall, A.G., Marcus, T. and Sallos, J. (1979) J. of Magn. Reson. 35, 227.

68. LePecq, J.B. and Paoletti, (1967) J. Mol. Biol. 27, 87.

69. Krugh, T.R., Nuss, M.E. in Biological Applications of Magnetic Resonance (Shulman, R.G., Ed.) pp 113-177, Academic Press, New York.

70. Tsai, C.C., Jain, S.C. and Sobell, H.M. (1977) J. Mol. Biol. 114, 301.

71. Jain, S.C., Tsai, C.C. and Sobell, H.M. (1977) J. Mol. Biol. 114, 317.

72. Krugh, T.R. and Reinhardt, C.G. (1975) J. Mol. Biol. 97, 133.

73. Lee, C.H. and Tinoco, I., Jr. (1978) Nature 274, 609.

74. Patel, D.J. and Shen, C. (1978) Proc. Nat. Acad. Sci. 75, 2553.

75. Waring, M.J. (1968) Nature 219, 1320.

76. Blake, A. and Peacocke, A.R. (1968) Biopolymers 6, 1225.

77. Krugh, T.R., Wittlin, F.N., Cramer, S.P. (1975) Biopolymers 14, 197.

78. Tanford, C. (1961) Physical Chemistry of Macromolecules, Wiley, NY

79. Houssier, C., Hardy, B. and Fredericq, E. (1974) Biopolymers 13, 1141.

80. Dalgleish, D.G., Peacocke, A.R., Fey, G. and Harvey, C. (1971) Biopolymers 10, 1853.

81. Aktipis, S. and Martz, W.W. (1974) Biochemistry 13, 112.

82. Davanloo, P. and Crothers, D.M. (1976) Biochem. 15, 5299.

83. Aktipis, S. and Kindelis, A. (1973) Biochem. 12, 1213.

84. LeBret, M., and Chalvet, O. (1977) J. Molecular Structure 37, 299.

85. Manning, G.W. (1969) J. Chem. Phys. 51, 924.

86. Le Pecq, J.B. (197) Methods of Biochemical Analysis 20 41.

87. Felsenfeld, G. and Hirsham, S. (1965) J. Mol. Biol. 13, 407.